

**Meningococcal genome dynamics:
conservation and variability in *Neisseria meningitidis***

Kristian Alfsnes

Thesis for the degree of Philosophiae Doctor (PhD)



Centre for Molecular Biology and Neuroscience (CMBN)

Department of Microbiology

Oslo University Hospital HF Rikshospitalet

University of Oslo, Norway

2012

© Kristian Alfsnes, 2012

*Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1419*

ISBN 978-82-8264-500-3

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinssen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Table of content

Acknowledgements	p. ii
Abbreviations	p. iii
List of papers	p. iv
1. Introduction	p. 1
1.1. History	p. 1
1.2. Epidemiology	p. 1
1.3. Virulence factors	p. 2
1.4. Mc genome-related biology	p. 5
1.4.1. Genome characteristics	p. 5
1.4.2. The DNA uptake sequence (DUS)	p. 6
1.4.3. Transcription terminators (TTs)	p. 8
1.4.4. Genome stability and instability	p. 9
1.4.5. Mutation	p. 10
1.4.6. Phase variation (PV)	p. 10
1.4.7. DNA repair pathways	p. 12
1.4.8. Horizontal gene transfer (HGT)	p. 14
1.4.9. Transformation	p. 14
2. Aims of the study	p. 17
3. Summary of results	p. 18
4. General discussion	p. 20
4.1. The search for the DUS-specific receptor and DNA binding proteins	p. 20
4.2. Genome instability generated by PV	p. 22
4.3. The nature of recombination in transformation	p. 24
4.4. Genome instability driving evolution	p. 25
4.5. Methodological considerations	p. 27
4.5.1. Identification of DUS- and DNA binding proteins	p. 27
4.5.2. Genome instability exhibited by PV	p. 31
4.5.3. Recombination events delineated by genome sequencing	p. 34
5. Future prospects	p. 36
6. References	p. 38

Acknowledgements

This work was carried out at the Institute of Microbiology, Centre for Molecular Biology and Neuroscience (CMBN), at Oslo University Hospital (Rikshospitalet) and University of Oslo, during the period of April 2009 to May 2012. Financial support was provided by the Research Council of Norway. Additional funding provided by the Oxford Norwegian Scholarship, Internationalization funds at the Medical Institute, University of Oslo, the Ella & Robert Wenzin's scholarship, the Henrik Homans Minde scholarship and the Per Rygh scholarship.

I wish to express my gratitude to my main supervisor Tone Tønjum, for giving me this opportunity to pursue my scientific hunger. I would also like to thank my co-supervisor, Ole Herman Ambur, without whom I would've been lost in the primordial soup. I would like to thank both supervisors for believing in my work, for guidance and support along the way.

In addition I would like to thank the other members of the group, especially Stephan Frye who has been instrumental in critical thinking and playing the devil's advocate. Big thanks to Håvard Homberset for extensive technical help, and to the rest of the group for helpful inputs and ideas. Thanks to my collaborators abroad, especially to Derek Hood and his group in Oxford for making my stay in his lab which was socially and scientifically rewarding, as well as to the graduate students at Wadham College. I would also like to thank my scientific peers at the institute, my symbiotic co-workers in the office and the lab. Thanks to Jorrit Enserink whom fibrous-neck made this work a lot more cheerful and to Vegard Eldholm for laughs and extracurricular activities.

Thanks to my friends here, there and everywhere, you've been supportive and offered valuable distractions from the scientific work when needed.

I would particularly like to thank my pride, lioness Gunhild and cub Nikolai, for your patience, love and support. Also, thanks to my sister, Katrine, for well seasoned perspective. I'm forever indebted to my parents whose great genes and environment led me to where I am now, thank you for your unconditional love and support.

Abbreviations

8oxoG	7,8-dihydro-8-oxo-2'-deoxyguanosine
BER	Base excision repair
CFU	Colony forming unit/s
CPS	Capsular polysaccharides
DUS	DNA uptake sequence
faPy	Formamidopyrimidine
Gc	Gonococcus/i
HGT	Horizontal gene transfer
IR	Inverted repeat
LFQ	Label-free quantity (from the MaxQuant analysis)
Mc	Meningococcus/i
MMR	Mismatch repair
MS-MS	Tandem mass-spectrometry
NER	Nucleotide excision repair
Nla	<i>Neisseria lactamica</i>
Nt	Nucleotides
PV	Phase variation
SNP	Single-nucleotide polymorphism
SPA	Solid-phase assay
TLS	Translesion synthesis
Tfp	Type IV pilus/i
TT	Transcriptional terminator
USS	Uptake signal sequence
Wt	Wildtype

Bold text in the introduction refers to chapters.

Bold text in the results and discussion refers to the papers.

List of papers

- I Benam A. V., Lang E., **Alfsnes K.**, Fleckenstein B., Rowe A. D., Hovland E., Ambur O. H., Frye S. A. and Tønjum T. (2011) “Structure-function relationships of the competence lipoprotein ComL and SSB in meningococcal transformation”. *Microbiology* 157:5, 1329-1342

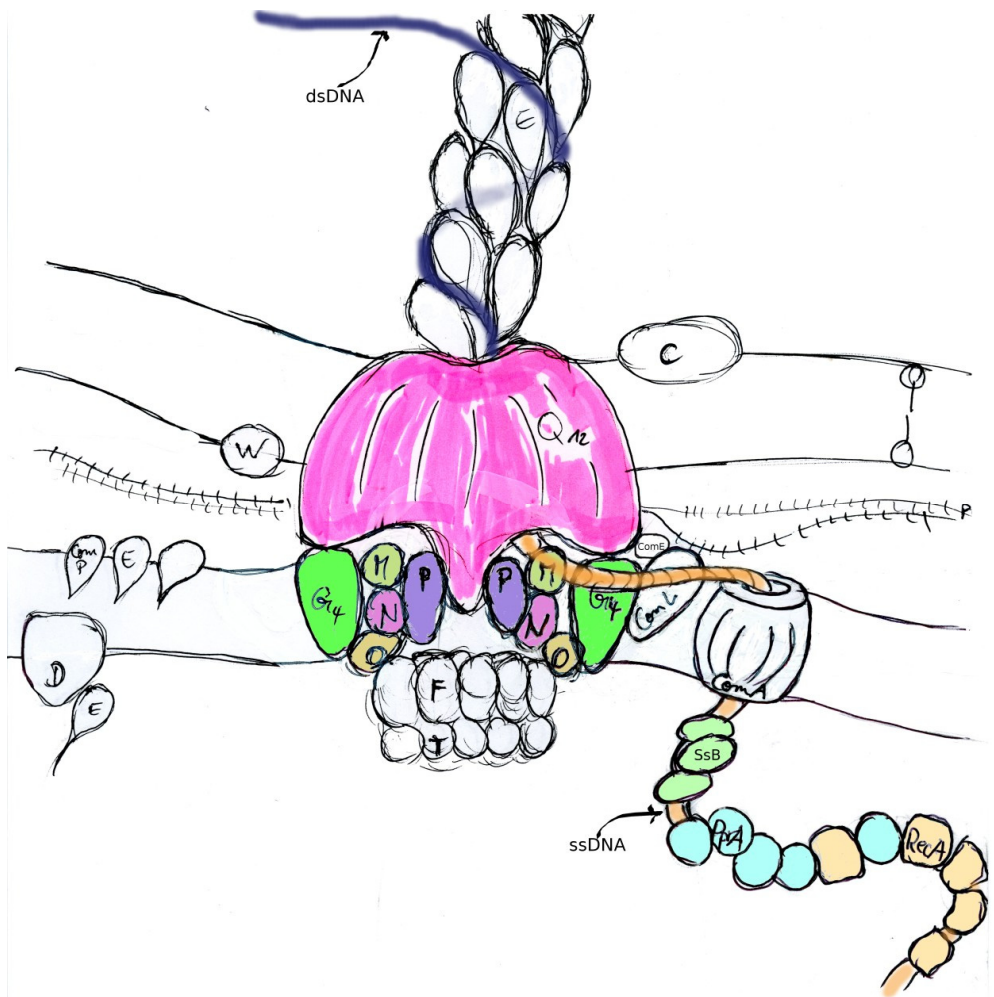
- II **K. Alfsnes**, W. Egge-Jacobsen, S. A. Frye, H. Homberset, T. Tønjum and Ambur O. H. “DNA uptake sequence-specific solid phase isolation of DNA binding proteins from *Neisseria meningitidis* cells”. (Manuscript)

- III Tibballs K., O. H. Ambur, **K. Alfsnes**, T. Davidsen, H. Homberset, S.A. Frye and Tønjum T. (2009) “Characterization of the meningococcal DNA glycosylase Fpg involved in base excision repair”. *BMC Microbiology* 9:7

- IV **K. Alfsnes**, T. Tønjum, and Ambur O. H. “The influence of homopolymeric G tract length and mismatch repair on phase variation in a *Neisseria meningitidis* model”. (Reviewed – pending rebuttal)

- V **K. Alfsnes**, O. Harrison, P. Power, S. Bentley, S. A. Frye, M. Maiden, D. Hood, T. Tønjum and Ambur O. H. “Intergenic crossover hotspots in meningococcal transformation”. (Submitted)

The papers will be referred to by their Roman numerals.



Drawing by SAF

*A man who dares to waste one hour of time has not
discovered the value of life.*

Charles Darwin

1. Introduction

1.1. History

Neisseria meningitidis, or the meningococcus (Mc), is a gram-negative member of the β -*Proteobacteria* subclass of bacteria. This feared bacterial pathogen, one of the main causative agents of meningitis and septicaemia, has been the focus of intense studies since the end of the 19th century. One of the early clinical descriptions of bacterial meningitis came during the 1805 epidemic in Geneva, Switzerland, which lasted 3 months and resulted in 33 deaths (Vieusseux, 1806). The first clinical isolation of the causative bacterium (then named *Diplokokkus intracellularis meningitidis*) succeeded in 1887 (Weichselbaum, 1887), while the important asymptomatic carrier state was discovered in 1896 (Kiefer, 1896).

Since then, a considerable scientific effort has been put into understanding the Mc lifestyle and pathogenesis and into the prevention and treatment of Mc disease. Despite these efforts, Mc is a leading causative agent of meningitis and septicaemia worldwide. The closely related *Neisseria gonorrhoeae*, the gonococcus (Gc), is localized in the human uro-genital tract. Gc is a sexually transmitted pathogen and has complemented *N. meningitidis* as a model organism in many cases. Mc and Gc have $\approx 96\%$ chromosomal sequence identity and are comparable human-specific organisms in many respects.

1.2. Epidemiology and relevance for vaccine development

Non-symptomatic Mc colonization takes place in the epithelial lining of the oropharyngeal mucosa in humans (DeVoe, 1982). Factors yet uncharacterized may allow Mc to enter the blood-stream and cross the blood brain barrier of the host, causing septicaemia and/or inflammation of the meninges and brain (Hardy *et al.*, 2000; Nassif, 2000).

Mc strains are classified by the antigenic properties of the capsular polysaccharide (CPS) (Gotschlich *et al.*, 1969) which defines their serogroup (Hammerschmidt *et al.*, 1996). Currently, a total of twelve Mc serogroups exist, six of which (ABCXYW₁₃₅) are responsible for 90 % of Mc invasive disease worldwide (Caugant, 2008). Development of vaccines based on high molecular weight cell-surface CPS can prevent infections of serogroup A, C, Y and W135 (Stephens, 2007). Although effective in adults, vaccines based on CPS do not sufficiently protect infants

and young children below the age of 2-5 years (who do not have fully developed immune systems) and individuals that do not induce immunological memory (Bilukha & Rosenstein, 2005; Stephens, 2007; Stephens *et al.*, 2007). A conjugate vaccine made by chemical linking Mc serogroup C CPS and a carrier protein has recently been shown to be protective in all age groups, inducing long-term immunological memory (Wing *et al.*, 2012). By the introduction of the Mc serogroup A vaccine “MenAfriVac”, the MenA epidemics in the Sahel region of sub-Saharan Africa (also termed the “meningitis belt”) is already greatly reduced and will according to predictions be completely eliminated (Roberts, 2010; Kristiansen *et al.*, 2011; Sow *et al.*, 2011). Right now the problem in Africa is the emerging of non-MenA epidemics, mainly due to serogroup W₁₃₅ (Boisier *et al.*, 2005; Forgor *et al.*, 2005).

For serogroup B, the CPS is poorly immunogenic in humans and is proven to be unfit due to bacterial mimicry of CPS of the human polysialated glycoproteins expressed in brain tissue (Finne *et al.*, 1983). Substantial efforts have been invested in alternative approaches, mainly with non-capsular antigens, outer membrane vesicles (OMV) (Bjune *et al.*, 1991; Cassio de Moraes *et al.*, 1992), natural vaccine using OMVs from *N. lactamica* (Nla) (Oliver *et al.*, 2002; Finney *et al.*, 2008; Evans *et al.*, 2011), and “reverse vaccinology” (Rappuoli, 2001; Giuliani *et al.*, 2006). Several serogroup B vaccines are currently being tested in clinical trials (Girard *et al.*, 2006; Sardiñas *et al.*, 2006; Jäkel *et al.*, 2008; Donnelly *et al.*, 2010). The Mc serogroup B disease (Caugant *et al.*, 1987) is still emerging on a worldwide basis, due to the lack of an effective vaccine against this serogroup, and MenB has caused epidemics in several countries including Norway, New Zealand, Brazil, Chile, Colombia and Cuba (Caugant, 1998; Tzeng & Stephens, 2000; Stephens, 2007; Stephens *et al.*, 2007). In the period 2001-2008, a strain specific vaccine was successfully used in New Zealand to control an epidemic outbreak of MenB meningitis (Arnold *et al.*, 2011).

1.3. Mc virulence factors

The ability of Mc to cause disease in its exclusive human host is correlated with adherence to mucosal epithelial cells in the oro- and nasopharynx and further invasion of subepithelial tissues and the blood stream, as well as the meninges, avoiding immune defences (Stephens *et al.*, 2007). Mc is predominantly found in the oropharynx, where it can be a constituent of the normal microbial flora. However, some strains are invasive, first penetrating into the blood stream and then the

meninges, to cause septicaemia and meningitis, respectively. The clinical signs and symptoms due to meningitis are the sum of the inflammation and brain oedema induced in the host, caused by Mc virulence factors such as lipooligosaccharide (LOS), porins and pili (Pujol *et al.*, 2000; van Deuren *et al.*, 2000; Brandtzaeg *et al.*, 2001; Brandtzaeg & van Deuren, 2002). The primary Mc virulence factors are the CPS and type IV pili (tfp), where the former prevents activation of the alternate complement pathway and the latter is required for adhesion and subsequent infection (Hammerschmidt *et al.*, 1996; Pujol *et al.*, 2000). IgA protease and factor H binding protein are other factors promoting survival and pathogenicity and are vaccine candidates under consideration (Seib *et al.*, 2009; Pajon *et al.*, 2010; Lucidarme *et al.*, 2011; Marsh *et al.*, 2011). The opacity proteins are involved in the adhesion to the host cells (Dehio *et al.*, 1998; Nassif *et al.*, 1999; Sadarangani *et al.*, 2011). Two outer membrane proteins (OMPs), the porins, PorA and PorB, occur abundantly and have been linked to virulence. PorA negatively regulates the complement response while PorB facilitates attachment and invasion of host cells (Rudel *et al.*, 1996; Massari *et al.*, 2000; Jarva *et al.*, 2005).

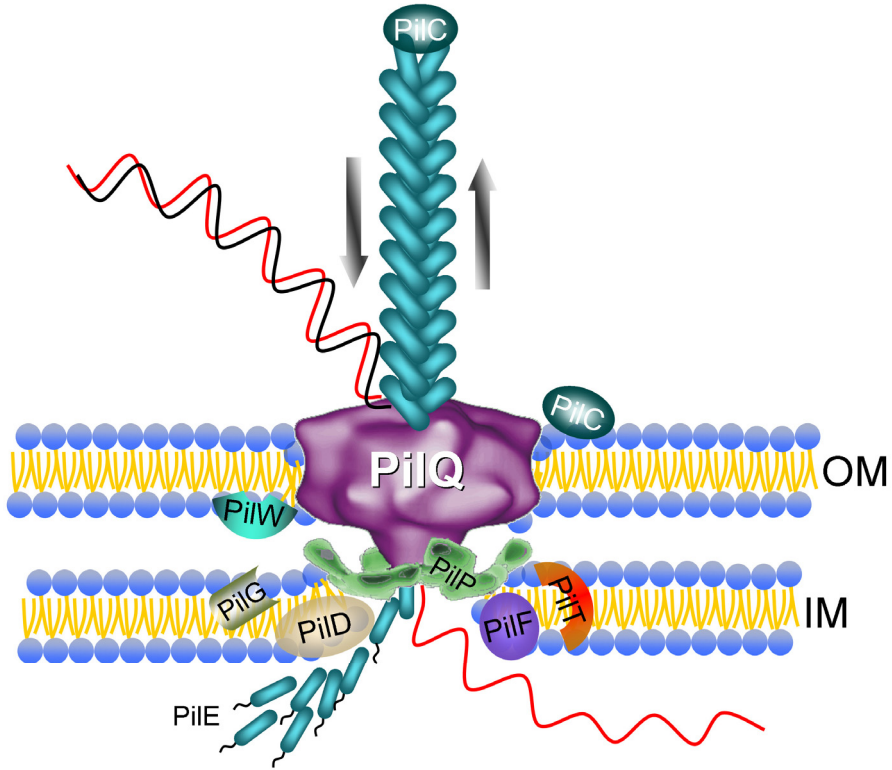


Fig. 1 - The type IV pilus (tfp) biogenesis complex and the transformation machinery are located mainly in the Mc membranes. OM = Outer membrane, PG = Peptidoglycan layer, IM = Inner membrane (made by SAF)

Tfp biogenesis is performed by a machinery (Fig. 1) which consists of the type II secretion system which itself branches from the general secretory pathway (GSP) (Pugsley, 1993). The GSP is broadly conserved among Gram-negative bacteria and allows the delivery of macromolecules and virulence factors across the membranes to the surroundings (Nunn, 1999; Pujol *et al.*, 2000; Vignon *et al.*, 2003; Johnson *et al.*, 2006). Tfp are involved in adhesion to host cells (Swanson, 1973), twitching motility (Swanson, 1978), antigenic variation and virulence (Koomey *et al.*, 1987; Nassif *et al.*, 1993). Tfp are also functionally linked to natural **transformation** of DNA (see section 1.4.9.).

1.4. Mc genome-related biology

1.4.1. Genome characteristics

The genomes of Mc strains Z2491 and MC58, representing serogroup A and B, respectively, were the first to be sequenced and annotated (Parkhill *et al.*, 2000; Tettelin *et al.*, 2000). These genomes and that of Mc serogroup C strain FAM18 (Bentley *et al.*, 2007) are complete and annotated. Recently, however, the genomes of a multitude of Mc and Gc clinical isolates and strains of other *Neisseriaceae* species have become available to the public (Chung *et al.*, 2008; Schoen *et al.*, 2008; Rusniok *et al.*, 2009; Bennett *et al.*, 2010; Joseph *et al.*, 2010; Budroni *et al.*, 2011; Chen *et al.*, 2011; Schoen *et al.*, 2011). In addition many not yet assembled or annotated shotgun sequences are available in public databases. The Mc genome lacks defined pathogenicity islands (Perrin *et al.*, 2002), but encompasses a few prophages (Klee *et al.*, 2000; Schoen *et al.*, 2008). Frequent **horizontal gene transfer** (HGT) and other recombination events have rendered its genome with a polyphyletic character that does not fit well with conventional classifications and has challenged Mc strain characterization. Mc strains can, however, be classified using multilocus sequence types (MLST) and genome sequencing (Maiden *et al.*, 1998; Jolley *et al.*, 2004; Caugant, 2008).

Mc genomes have a size of approximately 2.2-2.3 Mb and harbour approximately 2000 coding regions making up between 75-80% of the genome. The Mc genomes have a G+C content of 51-52 %, and are characterized by an abundance of repeat sequences and phase variable genes (Parkhill *et al.*, 2000; Tettelin *et al.*, 2000; Bentley *et al.*, 2007). The most abundant repeated sequence in Mc is the **DNA uptake sequence (DUS)**, which is discussed below (see section 1.4.2.). Corra elements, small insertion elements of ~100 bp, are also found in abundance in the Mc genome (more than 700 copies), probably modulating transcription and gene expression (Buisine *et al.*, 2002; Snyder *et al.*, 2009). The clustered, regularly interspaced, short palindromic repeats (CRISPR) are probably responsible for defence against foreign genetic elements in prokaryotes (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel *et al.*, 2005). In Mc, the CRISPR associated proteins (CRISPR-Cas) are represented by the type II CRISPR-Cas system, wherein the protein, Cas9, in addition to the ubiquitous proteins, Cas1 and Cas2, generate CRISPR-based RNA and cleaves the target DNA (Makarova *et al.*, 2011). Polynucleotide tracts consisting of repeated homo- or heteropolymers are subject to expansion and/or retraction and are found in

or near the promoter region or in the coding sequence of a multitude of Mc genes. These tracts are generating phase variation, predominantly influencing genes encoding for surface exposed proteins (Snyder *et al.*, 2001; Martin *et al.*, 2003; Moxon *et al.*, 2006). **Phase variation** is discussed in more detail below (see section 1.4.6.).

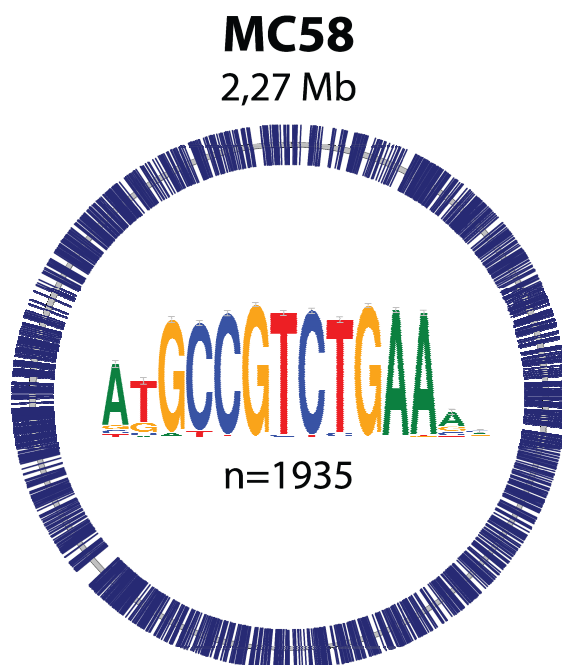


Fig. 2 – Depiction of the Mc strain MC58 genome exhibiting the presence and distribution of DNA Uptake Sequences (DUS) (represented as blue lines) (Frye *et al.*, unpublished).

1.4.2. The DNA uptake sequence (DUS)

The efficiency of transformation of Mc is greatly influenced by the presence of a short specific DNA sequence, termed the DNA uptake sequence (DUS) (Goodman & Scocca, 1988; Elkins *et al.*, 1991). Another specific genomic widespread sequence with a similar transformation increasing effect has been reported for *Haemophilus influenzae*, termed the uptake signal sequence (USS) (Sisco & Smith, 1979; Danner *et al.*, 1980). Transformation efficiency of Mc and *H. influenzae* are restricted to their species-specific sequences, DUS and USS respectively (Mathis & Scocca, 1982). The DUS in Gc was initially reported as a 10-mer sequence (Goodman & Scocca, 1988), however, it was later shown that the two nt longer and conserved 12-mer allowed for

even higher transformation rates (Ambur *et al.*, 2007). Extended work in our lab has recently demonstrated that different variants of the DUS sequence exist in closely related *Neisseria spp.*, termed DUS dialects (Frye *et al.*, manuscript submitted). Nearly 2000 copies of DUS are present in the Mc and Gc chromosomes (Fig. 2), constituting about 1% of the genome which makes DUS by far the most common 10-mer repeat in the Mc and Gc genome (Davidsen *et al.*, 2004) and can be said to actively sculpting the chromosome (Touzain *et al.*, 2011). The DUS are located both inside of genes (25%) and in intergenic locations (75%), the former suggestive of long term adaptive evolution due to the integration into the reading frame, the latter are often found as inverted repeat (IR) and suggestive of a **transcriptional termination** function (Ambur *et al.*, 2007). Our Genome dynamics and microbial pathogenesis group previously discovered a biased distribution of the intragenic DUS towards genome maintenance genes, and a preserving function of the genome by DUS-mediated transformation was proposed as a driving force behind the over-representation of DUS in coding sequences (Davidsen *et al.*, 2004). In favour of such a regenerative and conserving function, a refined analysis in a follow-up comparative genomics study demonstrated that DUS, by means of frequent recombination, had accumulated in regions of the conserved Mc, Gc and Nla core genomes and not in their accessory and variable regions (Treangen *et al.*, 2008). The DUS and the DUS-specificity therefore seem to have established itself during the evolution of the species driven by the positive effect that transformation and homologous recombination, and hence allelic reassortment, has on genome stability by maintenance of core functions. The DUS specificity varies between different Mc and Gc strains, ranging from low efficiency to obligatory for transformation to occur (Ambur *et al.*, 2007; Duffin & Seifert, 2010).

The hunt for the putative DUS-specific receptor, probably a protein, has been the focus of on-going studies over time (Chen & Gotschlich, 2001; Assalkhou *et al.*, 2007; Lång *et al.*, 2009); however, the identity of this hypothetical DUS-specific component remains elusive.

1.4.3. Transcription terminators

TTs in bacteria can be divided into two mechanistic groups, rho-dependent, requiring the rho protein or its homologs for termination, and intrinsic or rho-independent, initiated by dyad symmetry in the RNA transcript (Wilson & von Hippel, 1995). The rho-independent TTs have been shown to be more prevalent in *Mc*, while rho-dependent TTs are predominant in other species such as *Escherichia coli* (Banerjee *et al.*, 2006; Kingsford *et al.*, 2007). Inverted repeats in the genomic sequence may function as rho-independent transcription terminators (TTs) by forming stem-loop structures which mediate termination or attenuation of transcription (Wilson & von Hippel, 1995; Kingsford *et al.*, 2007). Such hairpins may destabilize the ternary transcription complex (Yager & Von Hippel, 1991) which causes RNA elongation to stop at the position of termination (Farnham & Platt, 1981) or destabilize specific binding between the polymerase and the nascent RNA (Altmann *et al.*, 1994).

In *Mc*, inverted repeat DUS (IR DUS) have been shown to form stem-loops (hairpin structures) immediately downstream of the termination codon, and are assumed to function as terminators of RNA polymerase (Goodman & Scocca, 1988). Approximately 50% of **DUS** are arranged as IR DUS and are likely to be functional rho-independent TTs or transcriptional modifiers (Goodman & Scocca, 1988; Ambur *et al.*, 2007; Kingsford *et al.*, 2007).

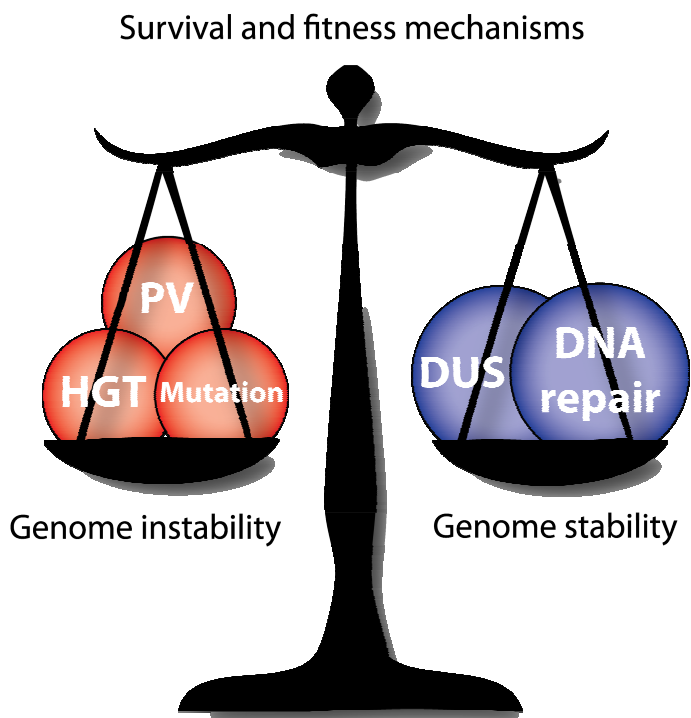


Fig. 3 - Schematic diagram depicting the balance between genetic variation and conservation in Mc.

1.4.4. Genome stability and instability

The genomes of all bacteria are constantly challenged in a balancing act to maintain and generate optimal conservation and variation (Fig. 3). The challenges are different for all bacteria with their respective niches / environments and life histories, and therefore adaptive mechanisms have evolved accordingly. Mc has developed an impressive array of mechanisms to balance genetic stability and instability.

One Mc survival strategy is to spontaneously produce a plethora of genomic variation (by **mutation** and **phase variation**) within the population to cope with changes in the environment rather than sensing changes (Moxon *et al.*, 1994; Feil *et al.*, 1999; Feil *et al.*, 2000; Feil & Spratt, 2001). As such, the genome is constantly under pressure for genetic change to produce an adaptive potential for the Mc cells, generating extensive genomic instability. The instability is caused by mutagens, mutations and slippage during replication and through **HGT** and recombination (Davidsen & Tønjum, 2006). In order to retain genomic integrity, Mc has evolved extensive **DNA repair pathways** to detect and correct mutations ensuring genomic

stability. Recombination can infer stability through **transformation** and recombination of homologous DNA containing the **DUS**, but also instability through **transformation** and recombination of sequence divergent DNA (Claverys *et al.*, 2009). In general, the balance between genetic variation and conservation defines the outcome in terms of adaptation and antigenicity, which in turn influence the microbial fitness for survival, virulence and drug-resistance (Tønjum & Seeberg, 2001).

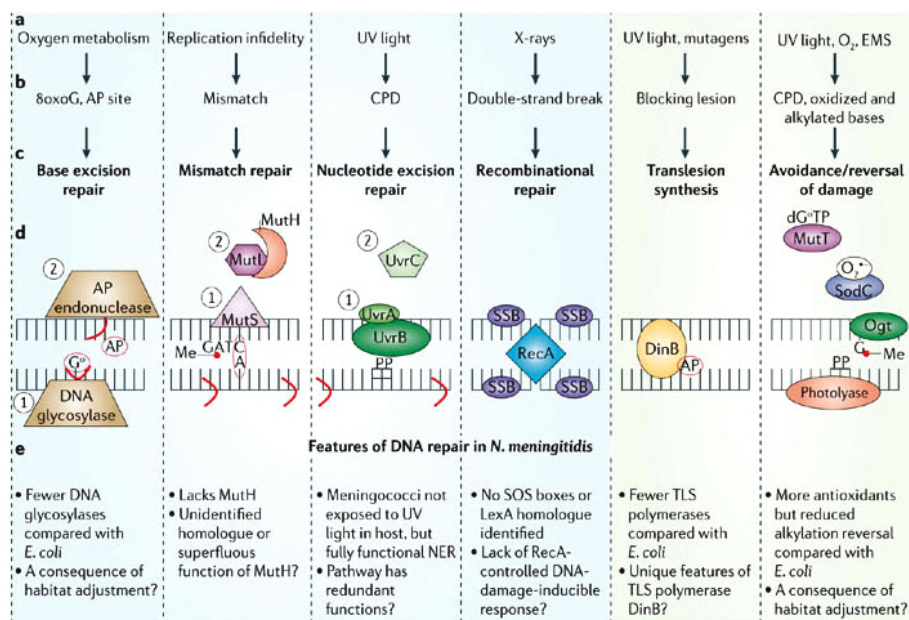
1.4.5. Mutation

Like in all other organisms, the Mc genome is subjected to a plethora of DNA damaging agents; external (environmental) and internal (metabolic) sources are causing spontaneous decay of the DNA in various ways (Lindahl, 1993). Oxidative stress can originate from both the environment (the human host) or from internal oxygen metabolism and is considered a major threat to genome integrity. Even though the Mc is protected from UV radiation in its habitat (the human host), it has a fully functional nucleotide excision repair (NER) pathway dealing with damages due to UV induced stress (Davidsen *et al.*, 2007; LeCuyer *et al.*, 2010). Replication infidelity in the single Mc polymerase DinB probably accounts for a large proportion of the Mc mutations, causing base-mispairing and insertion / deletion loops (Davidsen *et al.*, 2007). However, it has been shown that genetic changes in Mc occurs more frequently through DNA recombination than mutation, at a recombination to mutation ratio of at least 80 per-site within a Mc house-keeping gene (Feil *et al.*, 1999; Didelot & Falush, 2007). An MLST study of Mc, Gc and Nla suggested that the frequency of recombination between species is low (Bennett *et al.*, 2007), indicating that the majority of the genetic changes by DNA recombination occurs between strains and within distinct populations.

1.4.6. Phase variation (PV)

Phase variable genes, also termed contingency loci, are genes that can be switched on or off (reversible expression of antigens) and change the gene expression with simple nucleotide changes (Moxon *et al.*, 2006). The switch in expression occurs by changes in a polynucleotide tract found either in or near the promoter region, or in the open reading frame of the phase variable gene (Davidsen & Tønjum, 2006). The polynucleotide tracts are made up by homopolymeric or heteropolymeric repeat units that through slippage during DNA replication may gain or loose repeats and change

the reading frame (translation) or the promoter region (transcription) of the phase variable gene (Moxon *et al.*, 1994; Alexander *et al.*, 2004; Davidsen & Tønjum, 2006). The frequency of PV, switch in expression, of the repeated tracts are influenced by *cis*-acting factors (such as repeat unit, repeat length, repeat nucleotides) and by *trans*-acting factors (such as **DNA repair pathways** and DNA replication fidelity) (Bayliss *et al.*, 2001; Bayliss *et al.*, 2004). PV is a widespread mechanism, and has been found to influence the expression of genes in many bacteria (Henderson *et al.*, 1999; van der Woude & Bäumlér, 2004). Repeated sequences can also have translational importance to several human diseases affected by nucleotide expansion and retraction, such as Huntington and ataxia (Andrew *et al.*, 1993; Orr *et al.*, 1993). Mc hosts an abundance of phase variable genes, with varying estimates up to 100 different loci (Saunders *et al.*, 2000; Bayliss *et al.*, 2001; Snyder *et al.*, 2001; Martin *et al.*, 2003; Bentley *et al.*, 2007). Most phase variable genes encode components that are associated with the outer bacterial surface and therefore prone to be immunogenic, offering plasticity by a stochastic response to environmental stress, such as the host immune response (van der Woude & Bäumlér, 2004). PV may be considered a programmed event as the distribution of error-prone repeats is not random, making certain shifts in expression more likely to happen (Henderson *et al.*, 1999). It has been hypothesized that, by contributing to antigenic variation, phase variable genes enable Mc cells to evade the human immune system (Hammerschmidt *et al.*, 1996).



Copyright © 2006 Nature Publishing Group
Nature Reviews | Microbiology

Fig 4. The major Mc DNA repair pathways compared to *E. coli* (Davidsen & Tønjum, 2006)

1.4.7. DNA repair pathways

DNA repair pathways are distinguished by the nature of the damage they repair (Fig. 4) (Davidsen & Tønjum, 2006). Base excision repair (BER) is the main defence against base damages of endogenous origin especially oxidative stress (Seeberg *et al.*, 1995). Nucleotide excision repair (NER) repairs bulky lesions from exogenous sources, interfering with normal base-pairing and impairing transcription and replication (Nouspikel & Hanawalt, 2002). Mismatch repair (MMR) recognizes base-base mismatches and insertion/deletion loops introduced by the infidelity of DNA polymerases (Schofield & Hsieh, 2003). Recombinational repair (RR) handles double-strand breaks, and translesion synthesis (TLS) allows replication past blocking lesions (Napolitano *et al.*, 2000; Davidsen & Tønjum, 2006). Two of the pathways, MMR and BER have been shown to drastically affect the generation of mutator phenotypes in Mc (Oliver *et al.*, 2000; Richardson & Stojiljkovic, 2001; Richardson *et al.*, 2002; Macia *et al.*, 2005; Davidsen *et al.*, 2007), whereof MMR has also been shown to strongly affect the frequency of switching by **phase variation** (Richardson & Stojiljkovic, 2001; Richardson *et al.*, 2002).

The base excision DNA repair pathway (BER) is the main defense against the mutagenic and cytotoxic effects of endogenous lesions such as 7,8-dihydro-8-oxo-2'-deoxyguanosine (8oxoG) and formamidopyrimidine (faPy) (Seeberg *et al.*, 1995), and Mc cells have a fully developed BER pathway (Davidsen *et al.*, 2007). BER is initiated by a DNA glycosylase which identify and excise the damaged base, leaving an abasic (apurinic/apyrimidinic) site, which is then processed by the other enzymes of the BER pathway (Dempfle & Harrison, 1994; Davidsen & Tønjum, 2006). Mc cells express several DNA glycosylases, including Fpg (MutM), the major enzyme for removing 8oxoG from an 8oxoG:C base pair (Davidsen *et al.*, 2007), and MutY which releases 8oxoG from an 8oxoG:A mismatch (Li *et al.*, 2000; Li & Lu, 2003; Davidsen *et al.*, 2005).

MMR has been extensively investigated in *E. coli* and much of the current knowledge on MMR stems from this organism (Modrich & Lahue, 1996; Schofield & Hsieh, 2003). In short, the process of MMR can be summarized as follows: MutS recognize and binds the base mismatch and recruits the ATPase MutL, which orchestrates several interactions including the activation of MutH, an endonuclease nicking the un-methylated strand of newly synthesized DNA at GATC sites (Schofield & Hsieh, 2003). A piece of the nascent DNA that has received a nick is degraded by exonucleases with the aid of the DNA helicase UvrD before DNA polymerase III accurately re-synthesizes DNA, and the remaining nick is sealed by DNA ligase. The process also depends on single-strand binding protein (SSB) and the initial methylation of GATC sites by Dam methylase (Schofield & Hsieh, 2003). The strand containing the mismatch is degraded in the 5'-to-3' or 3'-to-5' direction, depending on the location of the mismatch relative to the nick (Iyer *et al.*, 2006). It is now clear, however, that all bacteria outside the γ -proteobacteria lack a MutH homolog (Ambur *et al.*, 2009). The endonuclease activity of MutH in *E. coli* and other γ -proteobacteria therefore seems to be constituted as an integral part of MutL in other bacteria such as Mc (Kadyrov *et al.*, 2006; Ambur *et al.*, 2009; Pillon *et al.*, 2010). It has recently been shown that the Mc DNA repair pathways are flexible in that Mc BER contributes to the repair of oxidative DNA damage interacting with MMR, so that MMR functions as a back-up for the BER system (Nagorska *et al.*, 2012).

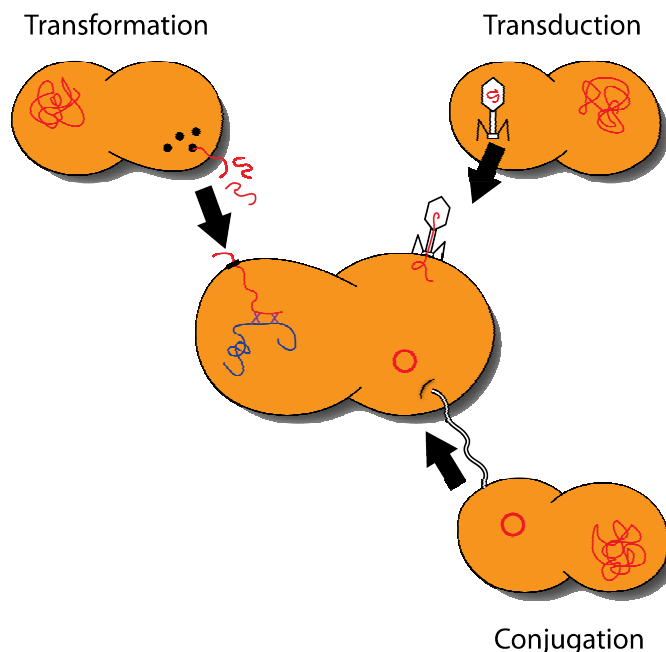


Fig. 5. The three modes of HGT in bacteria.

1.4.8. Horizontal gene transfer (HGT)

Horizontal gene transfer (HGT) allows for genes and genetic information to flow between organisms, to cross species barriers and beyond, and for the reshuffling of genes or the acquisition of novel genes not present in the available genetic pool. The selection for variation (antigenic variation, fitness for survival and virulence) is fuelled by a combination of recombination and mutation events as well as frequent allelic exchange by transformation which is one of three mechanisms for HGT. The three main HGT mechanisms are conjugation (plasmid-mediated), transduction (phage-induced) and transformation (uptake of naked DNA) (Fig. 5).

1.4.9. Transformation

Transformation, one of three modes of HGT in bacteria (Fig. 5), is defined by binding, uptake and genomic integration of exogenous DNA through recombination (Chen & Dubnau, 2003). Bacterial transformation is comparable to sexual reproduction in eukaryotes; they both involve active translocation of DNA from one cell to another, as well as homologous recombination. The transformation process in many bacteria is regulated and may require stress-induction, whereas in *Mc*,

transformation is constitutive (Lorenz & Wackernagel, 1994; Koomey, 1998). Recent studies indicate that natural transformation is more widespread in bacteria than previously thought and may have evolved several times independently (Johnsberg *et al.*, 2007; Claverys *et al.*, 2009). For example the uptake pathways in Mc and *H. influenzae* are closely related, and both species have evolved their own specific sequence for increased transformation efficiency, namely **DUS** and USS, respectively (Frye *et al.*, manuscript submitted) (Chen & Dubnau, 2004). In contrast to Mc, the natural competent bacteria *H. influenzae*, *Bacillus subtilis* and *Streptococcus pneumoniae* have not been shown to be competent through-out their life-cycle (Jyssum & Lie, 1965; Singh & Pitale, 1968; Håvarstein *et al.*, 1995; Macfadyen *et al.*, 2001).

Transformation is thought to be the predominant source of new genetic information in *Neisseria sp.* (Fussenegger *et al.*, 1997; Koomey, 1998; Davidsen & Tønjum, 2006). Mc and Gc are naturally transformable throughout their life-cycle (Jyssum & Lie, 1965; Sparling, 1966). Prerequisites for the neisserial transformation process are the expression of *tfp* (Fig. 1), the presence of **DUS** in the incoming DNA (Fig. 2) and RecA-mediated homologous recombination (Biswas *et al.*, 1977; Mathis & Scocca, 1982; Koomey & Falkow, 1987; Goodman & Scocca, 1988). The *tfp* structures required for transformation include the pilin structural subunit PilE and the minor pilin ComP (Wolfgang *et al.*, 1999), as well as the outer membrane secretin PilQ which forms a pore through which pili are extruded and retracted (Tønjum *et al.*, 1998; Collins *et al.*, 2005; Frye *et al.*, 2006). Interestingly, PilV has been shown to function antagonistically to ComP, i.e. reducing DUS-specific uptake of DNA (Aas *et al.*, 2002). Inner membrane proteins include; PilG, PilP and PilT, for pilus assembly, secretin stabilization and pilus retraction respectively (Tønjum *et al.*, 1995; Wolfgang *et al.*, 1998; Balasingham *et al.*, 2007). We hypothesize that transforming DNA is transported into the cytoplasm in the wake of the retracting pili through the secretin pore, supported by the associated *tfp* proteins (Davidsen & Tønjum, 2006; Assalkhou *et al.*, 2007). Once DNA is inside of the outer membrane of the Mc cell, non-pilus associated proteins ensures transport through the periplasm, the peptidoglycan layer supported by ComL and Tpc and through the cytoplasmic membrane facilitated by ComE binding (Fussenegger *et al.*, 1996a; Fussenegger *et al.*, 1996b; Fussenegger *et al.*, 1997; Chen & Gotschlich, 2001; Lång *et al.*, 2009; Benam *et al.*, 2011), finally the DNA is assisted through the cytoplasmic membrane by the transmembrane channel

ComA (Facijs & Meyer, 1993; Claverys & Martin, 2003). Restriction modification enzymes and the helicase activity of the RecBCD pathway process transformed dsDNA to ssDNA in the cytoplasm (Mehr & Seifert, 1998). In the terminal steps of the transformation process, the cytoplasmic protein SSB protects the DNA (Chen *et al.*, 2005), followed by recruitment of DprA and RecA enabling homologous recombination (Koomey & Falkow, 1987; Bergé *et al.*, 2003). In *H. influenzae*, homologous recombination has been shown to be modulated by topoisomerase A (TopA) (Chandler & Smith, 1996).

These insights into the life-history and genetics of Mc form the basis upon which this thesis builds; the dynamic genome of Mc, modulated by DNA repair pathways, mutations and phase variations, and the DUS-specific DNA transformation.

2. Aims of the study

The main aim of this thesis was to characterize meningococcal genome plasticity, with emphasis on drivers of genome variation and conservation, through mutations, natural transformation and DNA repair.

The distinct sub-goals for this study were the following:

1. To identify and characterize meningococcal DNA binding proteins.
2. To search for the neisserial DUS-specific DNA binding proteins.
3. To define the role of defects in meningococcal Fpg on mutability.
4. To define the effect of polyG tract lengths on phase variation and colony size in wildtype and MMR mutant backgrounds.
5. To define the extent and limits of homologous recombination in transformation.

3. Summary of results

The mechanisms responsible for transformation and adaptation in *Mc* are not fully understood. **Papers I** and **II** of this project aimed to identify novel DNA binding proteins and proteins involved in the multi-factorial transformation machinery. In **Paper I**, a solid-phase overlay assay was executed with DNA substrates in combination with 1D and 2D gel electrophoresis to reveal potential DNA binding proteins from membrane and soluble *Mc* cellular fractions. The proteins were excised from the gel and submitted for peptide mass fingerprinting by MALDI-TOF mass-spectrometry for identification. A single protein, SSB, was identified from the soluble fraction, and five proteins from the inner membrane fraction. Two of these inner membrane proteins, PilG and PilQ, constitute part of the Tfp. The proteins identified bound non-specifically to ssDNA and dsDNA, but revealed no DUS specificity. Null-mutants of the two hypothetical proteins identified (NMB1796 and NMB1963) from the inner membrane fraction, exhibited no change in competence for transformation. Null-mutants corresponding to the predominant protein from the inner membrane fraction, ComL, and the protein in the soluble fraction, SSB, could not be made since they are essential genes. In **Paper II**, isolation of DNA- and DUS-specific proteins in solution was obtained by immobilization of proteins from cellular fractions by binding to biotinylated ssDNA and dsDNA probes with and without DUS bound to streptavidin-covered magnetic beads. The immobilized hypothetical DNA- and DUS-specific binding proteins were identified using tandem mass-spectrometry (MS-MS) and quantified using the MaxQuant software. The quantity, measured as label-free quantity (LFQ) intensity with MaxQuant, of proteins identified using random DNA- or DUS-containing probes were compared to each other and to the corresponding proteins detected with the probe-less negative control. A scrutinizing threshold was adopted from a similar study, resulting in a small panel of potential DNA and DUS-specific binding proteins. One candidate protein, TopA, showed a significantly elevated affinity to the DUS-containing DNA probes over the random DNA probes.

The expressions of many surface proteins in the *Neisseria* are known to vary due to polynucleotide instability, a process known as phase variation (PV). PV represents a simple, efficient mechanism to quantify polynucleotide-instability and hence genome instability *in vivo* and was employed to study genetic instability in a quantitative manner (**Papers III** and **IV**). The study presented in **Paper III** focused

on the effect of the BER DNA glycosylase Fpg, compared to the MMR protein MutS, in a PV assay. PV frequency for a contingency locus in the *fpg* null mutant was, compared to the wildtype (wt) and a *mutS* null mutant. The *fpg* mutant exhibited an intermediate increase (2-fold) in PV frequency compared to the wt, comparably less than the mutator effect observed in the *mutS* null mutant (30-fold).

In **Paper IV**, the PV frequencies of a panel of polyG tracts with lengths of 5-13 nucleotides (nt) were studied in the wt and MMR mutant (*mutL*) backgrounds. In addition to quantifying the frequency of OFF→ON switching of the phase variable gene inferring antibiotic resistance, measurements of the colony size of the panel of polyG tracts in both backgrounds were made to provide an estimate of the level of back-switching (ON→OFF). PV frequency was found to be strongly affected by both the tract length and MMR; the highest PV frequency was observed for intermediate tract lengths (8-11 nt), while the MMR mutant showed higher frequency in the lower range of tract lengths and generally higher frequency for all tract lengths compared to the wt background. The MMR mutant was found to grow to smaller colony sizes than the wt background for short tracts; the sizes were found to converge and were reduced in both backgrounds as the tract lengths extended. Based on the reduced colony size for the longer tracts, we proposed that back-switching (ON→OFF) occurs more frequently when the tract length increases.

Cross-over hot-spots (COHs) from homologous recombination in transformation were studied in **Paper V**, where a Mc strain was selectively transformed with donor DNA from another Mc strain conferring antibiotic resistance through a point mutation. The complete genomes of pooled transformants were sequenced using high-throughput sequencing, revealing the genome-wide extent of homologous recombination events by comparing the single-nucleotide polymorphism (SNP) profiles between the two original strains and the resulting transformed Mc. The COHs were located in regions of sequence divergence between the recipient and donor DNA. The COHs were also found to coincide with TTs, some of which contained DUS IRs in their stem.

Taken together, the studies in **Papers I-V** contribute to elucidating the two evolutionary forces sculpting the Mc genome; conservation and variation of DNA.

4. General discussion

This thesis focuses on how Mc cells allow extensive genetic variation while securing genetic conservation (Fig. 3). Genome instability can be generated by mutation, recombination and horizontal gene transfer (HGT) (such as transformation). Mc cells readily take up and incorporate extracellular DNA into their genome through transformation, providing new coding ability (Jyssum & Lie, 1965; Davidsen & Tønjum, 2006). Genetic variation introduced by transforming DNA contributes to genetic variation and thereby increases the adaptive potential within populations of Mc. At the same time, the DUS-specific Mc transformation machinery mediates conservation by scrutinizing the accessible DNA for the presence of DUS, ensuring transformation of homologous DNA (Davidsen *et al.*, 2004; Treangen *et al.*, 2008). The studies of non-specific and DUS-specific DNA binding proteins contribute to illustrating how parts of this process occur at the protein level (**Paper I and II**). Characterizations of Fpg, a DNA glycosylase of the Mc BER, revealed how genetic conservation is secured through this mechanism, but also how defects in Fpg modulate PV and cause variation (**Paper III**). PV was also used to study how variation is modulated by the genetic structure by introducing a range of phase variable tract lengths and how this variation is influenced by a defect in MMR (**Paper IV**). Cross-over hotspots of recombination revealed how divergent genetic variability in intergenic regions disrupted recombination of these regions between different Mc strains, conserving intragenic regions through homologous recombination, while allowing variation in non-recombined intergenic regions (**Paper V**). Taken together, this thesis aimed to enlighten Mc genome dynamics by delineating new aspects of the balancing act between genetic variation and conservation.

4.1. The search for the DUS specific receptor and DNA binding proteins

The elusive hypothetical DUS-specific binding protein(s) in neisserial transformation have challenged scientists since the discovery of the specificity in neisserial transformation (Graves *et al.*, 1982; Goodman & Scocca, 1988). In an effort to characterize novel DNA-binding (and potential DUS-specific) proteins, 1D and 2D electrophoresis was used in combination with a solid-phase overlay assay and subsequent MS analysis (**Paper I**). The identification of the membrane-associated DNA binding proteins, ComL and PilG, confirmed previous studies showing DNA

binding affinity of these proteins (Fussenegger *et al.*, 1996a; Lång *et al.*, 2009). The identification of the DNA binding SSB protein in the soluble fraction confirmed previous studies of the influence of the abundant SSB in processing of ssDNA during DNA replication, repair and recombination in *E. coli* (Shereda *et al.*, 2008). The SSB homolog in *S. pneumoniae* has been shown to influence competence (Campbell *et al.*, 1998), suggestively by binding to the 3' donor DNA, forming a DNA-RecA filament before recombination (Dubnau, 1999). The non-viability of the *comL* and *ssb* mutants indicates that these proteins are essential for neisserial survival. Studies in other organisms have shown that homologs of ComL in Gc (Fussenegger *et al.*, 1996a), and homologs of SSB in *B. subtilis* (Lindner *et al.*, 2004) and in *E. coli* (Shereda *et al.*, 2008) are indeed essential for survival. The identification of two hypothetical proteins, for which the null-mutants retained competence, was assumed not to be associated with DNA binding / processing in transformation. PilG and PilQ have previously been shown be involved in tfp biogenesis (Tønjum *et al.*, 1995; Tønjum *et al.*, 1998), indirectly affecting competence. The identification of the DNA binding ability of PilG and PilQ (Frye *et al.*, 2006; Lång *et al.*, 2009) presents the question whether these proteins also affect transformation directly. Notably, a new functional DNA binding fold has been identified in PilG (Lång *et al.*, manuscript submitted).

The second study combined a well-established method of protein immobilization with protein identification in a new way, utilizing pull-down of proteins with magnetic beads and subsequent MS-MS-based identification and quantification of proteins isolated (**Paper II**). Comparing the DNA probes against the probe-less control revealed seven proteins which ratified the threshold used by Målen *et al.* (Målen *et al.*, 2011); TopA, PolA, Fhs, GuaB, LpdA1, GdhA and SfcA. Two of these proteins, based on their functionality, have the potential to be DUS-specific and/or DNA-binding, namely TopA and PolA. The former candidate was found with relatively high LFQ intensity using all probes, while the latter candidate was only found using the dsDNA probes. Notably, the DNA topoisomerase A, TopA, exhibited a higher relative affinity when comparing DUS-containing DNA probes with the random DNA probes and the probe-less control. The TopA protein is a well-recognized DNA binding protein (Champoux, 2001) and is probably required for competence in *H. influenzae* (Chandler & Smith, 1996). Importantly, a hypothetical DUS-specific affinity of TopA to DNA has not previously been detected. Interestingly, the relative difference in affinity was found to be higher for ssDNA

probes with a free 5' end (biotinylated 3') than for ssDNA probes with free 3' end (biotinylated 5'), thus, Mc TopA may recognize single-stranded DUS in a 3'-5' direction in the DNA.

4.2. Genome instability generated by PV

The activity and specificity of recombinant Mc Fpg purified to homogeneity were examined with representative DNA substrates known to be induced by oxidative DNA damage, 8oxoG and faPy (**Paper III**). The substrate specificity assays confirmed Mc Fpg as a prototypical Fpg glycosylase protein, confirming substrate affinity observed in other model organisms, such as in *E. coli* (Tchou *et al.*, 1991). Since Mc Fpg effectively removes oxidized guanines, we asked the question if the protein also modulates PV of a repeated guanine (polyG) tract. Consequently, a PV assay was established to detect any mutator activity of the *fpg* null mutant as compared to the wt. An MMR mutant, *mutS*, used as a positive control, confirmed previous findings displaying a 30-fold increase compared to the background (Richardson & Stojiljkovic, 2001; Richardson *et al.*, 2002; Davidsen *et al.*, 2007). The Mc *fpg* mutant, however, only showed a modest, 2-fold increase, but nonetheless statistically significantly higher PV than the background activity. In a spontaneous mutation frequency assay, the Mc *fpg* single mutant exhibited only a weak mutator phenotype. However, Mc *fpg* mutant in concurrence with another BER mutant, *mutY* has been shown to exhibit a strikingly high-frequency mutator phenotype (Davidsen *et al.*, 2007). As such Mc Fpg *per se* seems of minor importance in modulating the organism's mutation frequencies, probably interacting with other BER proteins in a close interplay providing back-up between DNA glycosylases (Davidsen & Tønjum, 2006).

Paper IV builds on previous studies using a similar phase variable construct (Alexander *et al.*, 2004) and **Paper III**), as well as previous studies comparing the mutator ability of Mc DNA repair pathways (Richardson & Stojiljkovic, 2001). In this study we manipulated the length of the nucleotide repeat tract responsible for the PV, comparing the PV frequency of the various lengths in addition to the effect of a Mc MMR mutant background. Previous PV assays have observed and compared the phenotypic switching frequency in different PV loci (Richardson & Stojiljkovic, 2001), in clinical strains (Richardson *et al.*, 2002) and with different DNA repair pathway mutant backgrounds (Richardson & Stojiljkovic, 2001; Alexander *et al.*, 2004; Martin *et al.*, 2004). However, no previous study has assessed the impact of

different polyG tract lengths in wt and in *mutL* null-mutant background. Thus, **Paper IV** enabled the comparison of the influence of both a *cis*-acting effect (tract length) and a *trans*-acting effect (MMR mutant). It has previously been shown that MMR mutants drastically increased PV frequency in Mc (Richardson & Stojiljkovic, 2001; Bayliss *et al.*, 2004). Increased PV frequency with increased tract length has been reported for a tetranucleotide repeat tract in *H. influenzae* (De Bolle *et al.*, 2000), as well as for mononucleotide and trinucleotide repeats in yeast (Tran *et al.*, 1997; Freudenreich *et al.*, 1998). This study showed a non-linear increase in PV frequency for the panel of tract lengths in both the MMR mutant and wt background, where the highest frequencies were observed for the intermediate tract lengths. Additionally, this study showed that the MMR mutants increased PV frequency for the whole panel of tract lengths compared to the wt background. In the tract length assessment, an estimation of back-switching (loss of antibiotic resistance), was included to validate the observed non-linearity between PV frequency and tract length. Estimates of back-switching were obtained by calculating the colony size, measured as area (mm²). It was hypothesized that the instability increases linearly with tract length, but that the proportion of back-switching (to antibiotic sensitivity, Spc^s) occurs at an increasing frequency as the instability increases, thus confounding the expected linear increase in PV frequency. As the Spc^r founding colony grows, we hypothesized that for every division and for every Mc cell, there is a chance of reverting to the Spc^s state, thus reducing the observed growth. The colony size observations supported this hypothesis, revealing reduced colony size relative to the increasing tract lengths, indicative of instable Spc^r founding colonies. Additionally, the phase variable gene was sequenced pre- and post-PV as a control of the polyG tract change. The pre-PV sequences confirmed that all tracts were of the assigned length and indeed antibiotic-sensitive (Spc^s). The post-PV sequences confirmed that all colonies were antibiotic-resistant (Spc^r) and had gained or lost a single nucleotide, according to the closest in-frame mutation. Closer inspection of the sequences from the longest repeat tracts revealed so called stutter bands, generated by polymerase infidelity during PCR or sequencing of long homopolymeric tracts or repeats (Streisinger & Owen, 1985; Shinde *et al.*, 2003). This is a novel way of monitoring PV, combining two assays to observe both the acquisition and loss of the expression of an antibiotic resistance. Whereas previous estimates of ON→OFF switching only stem from models (Drake, 1991; Saunders *et al.*, 2003), this study offer this information through direct

observations of the consequence of tract instability in Mc. As detected in the colony size assay, back-switching measured as the reduction of colony size is increased from the shortest tract to the longest. Both assays observe the continuous back-switching, the early back-switchers do not produce viable colonies and will thus not be detected in either the colony size assay or in the PV assay. It is conceivable that the divergence from the linear increase in PV with increased tract length could be due to these hypothetical early back-switchers, colonies with unstable tracts that rapidly switch ON→OFF and become non-viable. The best model to explain the observed colony size differences were continuous back-switching during cell replication.

The highest PV frequency was observed for the intermediate tract lengths, 8-12, corresponding to the average homopolymeric tract lengths of all putative phase variable genes known in Mc. This suggests that due to the mechanistic of the PV, it seems evolutionary beneficial to retain a tract of intermediate length to allow for change, but avoid excessive back-switching.

The PV assay offers a simple way of measuring genetic change and instability of such contingency loci and has been extensively used and studied in Mc (Moxon *et al.*, 1994; Martin *et al.*, 2003; Moxon *et al.*, 2006). The widespread extent of phase variable loci in Mc indicates that the phenotypic expressions of many genes are modulated by PV frequency (Saunders *et al.*, 2000; Bayliss *et al.*, 2001; Snyder *et al.*, 2001; Martin *et al.*, 2003; Bentley *et al.*, 2007). A review included in **Paper IV**, of previously published phase variable loci were compiled in a list of 101 putative phase variable genes, with the majority represented with homopolymeric intragenic tracts. The average homopolymeric tract length in the three Mc strains investigated ranged from 8 to 9 nt.

4.3. The nature of recombination in transformation

Whole genome sequencing of a pool of Mc transformed with non-isogenic Mc donor DNA containing a single nucleotide mutation conferring antibiotic resistance allowed for a genome-wide snap-shot of a transformation event (**Paper V**). The analyses of the SNPs acquired in the transformed Mc revealed possible recombination of individual DNA fragments up to 15 kb in size in both directions of the transformed single nucleotide mutation. A correlation was discovered between the COHs and sequence divergence between the two strains (donor and recipient); the sequence divergence was strongest in the intergenic regions as were the COHs. Secondly, many

of these COHs and sequence divergence correlated with the presence of TTs, some of which are consisting of DUS arranged as IRs. It has previously been postulated that sequence divergence may account for COHs, but this is to our knowledge the first study to observe this notion using transformation and subsequent whole genome sequencing. The homology requirements of RecA in *E. coli* has been studied extensively and found the frequency of recombination to be dramatically decreased by sequence heterology (Watt *et al.*, 1985). Also, stretches of non-homologous DNA may also be discriminated against following the flap-endonuclease activity of RusA acting on the migrating D-loop (Rosenberg S., personal communication). Additionally, the IRs may halt the D-loop progression and hence the homologous recombination process as they form stem-loop structures could physically interfere with the D-loop. Studying a Mc transformation event on the genomic scale in this manner is the first of its kind.

4.4. Genome instability driving evolution

In all domains of life, the survival and propagation of a species depends on the adaptive potential, the ability to respond to a changing environment and to simultaneously retain genomic integrity. The scale of the environmental change vary according to the species in question, from the microscopic scale of host cells for virus and bacteria (Bohannon & Lenski, 2000), to the global climatic change affecting the survival of whales, the world's behemoths (Walther *et al.*, 2002). In some cases of near constant environment and in the cases of some generalists, genetic change can be minimal, shaping so called "molecular living fossils" (e.g. species of tree ferns (Soltis *et al.*, 2002) and possibly animals around hydrothermal vents (Little & Vrijenhoek, 2003)). At the microscopic scale, a human-specific bacterium, such as Mc, can be exposed to sudden burst of antibiotics or mutagens, whereby rapid genetic alterations (such as PV) or other frequent genetic variability (shuffling of alleles in a population by transformation, inversion and gene conversion) may generate adaptive traits, however, without the means of controlling the genetic variability, genomes may degenerate.

Bacteria can respond to environmental changes by two modes; by changing its gene expression (a change in phenotype, without changing the genotype) and by changes in the genotype. For bacteria, changes in environment such as antibiotic treatment may not impose the necessary means of response or detection. Genotype

changes, on the other hand, takes advantage of the average genome mutations (μ g) to respond to changes in the environment. However, since most mutations are detrimental, this adaptive strategy requires a functional maintenance and innate repair system (Rosenberg, 2001; Moxon *et al.*, 2006).

The level at which selective pressures act on bacteria varies from the level of populations, organisms, cells or genes/molecules, depending on the life history of the bacterium in question. Selection also simultaneously acts on several orders and levels, e.g. selecting for conservation of single genes at the same time as selecting for population diversity (Moxon *et al.*, 2006). Consequently, selection acts on different orders for a clonal bacterium (low HGT) than for a highly competent bacterium (high HGT). Transformation of naturally competent bacteria with exogenous DNA is an event like eukaryotic sex caused by recombination and hence allelic reassortment. When applying evolutionary theories of sex to bacteria, the level of natural selection can vary from changes in a single nucleotide, genes (or regions) to whole gene complexes. The various selective pressures differ among the species of bacteria, depending on their life-style and habitat. The evolutionary history behind the acquisition and genome-wide establishment of DUS in Mc is interesting in this respect as they are abundantly present in high frequencies in conserved regions in the genome, while they are absent in genes encoding surface components and in regions of phage origin (or other invasive elements) (Treangen *et al.*, 2008). It can be hypothesized that the DUS elements contribute through homologous recombination to conserve critical genome regulating regions (e.g. DNA repair and “housekeeping genes”), while phase variable genes ensure antigenic variation (Moxon *et al.*, 1994; Spratt *et al.*, 2001). Post-translational modifications are dynamic changes in the expression of proteins, such as modifications of the pilin subunit which discourage aggregative behaviour, and may allow for colonization of new sites or hosts (Aas *et al.*, 2007; Vik *et al.*, 2009; Borud *et al.*, 2010; Chamot-Rooke *et al.*, 2011). Additional variation is generated in Mc by invasive elements and non-functional DNA, such as transposons or phages, which are constantly removed by homologous recombination following promiscuous natural transformation (Lawrence *et al.*, 2001). Such variation has a major impact on virulence, antibiotic resistance, antigenicity and the population structure of bacterial pathogens. As such the Mc genome is self-regulatory, allowing for variation while ensuring conservation.

The various Mc strains differ in their natural transformation potential, preference in DUS recognition and spontaneous mutagenicity (Richardson & Stojiljkovic, 2001; Richardson *et al.*, 2002; Ambur *et al.*, 2007; Duffin & Seifert, 2010). Thus, DUS, in addition to its functional constraint in DNA transformation, may also modulate the transcription of genes. Another compelling hypothesis is that these intergenic IR DUS may halt or arrest homologous recombination by interfering with the Holliday junction complex, imposing a structure resolution. The various strains is thus employing different survival strategies by balancing genome instability vs. stability to allow for adaptation to a variable and hostile environment while retaining the most important cata- and metabolic pathways, functions and species identity. This constant spontaneous drive for variation in the pathogenic bacteria Mc to cope with the hosts' response is analogous to evolutionary arms-race and the Red Queen hypothesis seen in all domains of life (van Valen, 1973; Dawkins & Krebs, 1979).

4.5. Methodological considerations

Below follows a brief summary of the methodological constraints and limitations, as well as efforts to overcome these shortcomings in **Papers I-V**.

4.5.1. Identification of DUS- and DNA binding proteins

Two approaches were used and included in this thesis searching for DNA- and DUS-specific binding proteins; 1D and 2D electrophoresis and solid-phase overlay assay (South-Western) (**Paper I**), and a solid-phase assay (SPA) (Fig. 6) using biotinylated DNA probes attached to streptavidin coated paramagnetic beads (**Paper II**).

The search for DNA- and DUS-specific binding proteins using the 1D and 2D electrophoresis and solid-phase overlay assay (not to be confused with the solid-phase assay – see below) revealed putative DNA-binding proteins, but no protein was identified exclusively specific to the DUS substrate (**Paper I**). Enrichment of the Mc cellular fractions (soluble and membrane) decreased the target complexity, as opposed to whole-cell extracts. The 2D electrophoresis of the solid-phase overlay assay increased the spatial resolution of the hypothetical DNA-binding proteins compared to 1D electrophoresis. Two of the DNA-binding proteins observed were found to be non-essential for competence, possibly engaging in other parts of DNA metabolism than

transformation. Two proteins were found to be essential for viability, and finally two proteins were found to be essential for competence. The absence of DUS-specific binding proteins can be due to technical limitations, in that DUS-specific binding can be very transient with variable affinity and in that the *in vitro* conditions in the solid-phase overlay assay do not truly enough reflect how DNA binding conditions *in vivo*. The solid-phase overlay assay is limited by the level of protein expression and by protein renaturation abilities after SDS-PAGE. However, the reproducible identification of DNA binding proteins (Lång *et al.*, 2009) and congruent results to other independent assays (Assalkhou *et al.*, 2007) represent strong evidence that the method is indeed a technique for detecting DNA binding proteins in general. However, the next logical step following the limitations of solid-phase overlay assay was the liquid phase immobilization (**Paper II**).

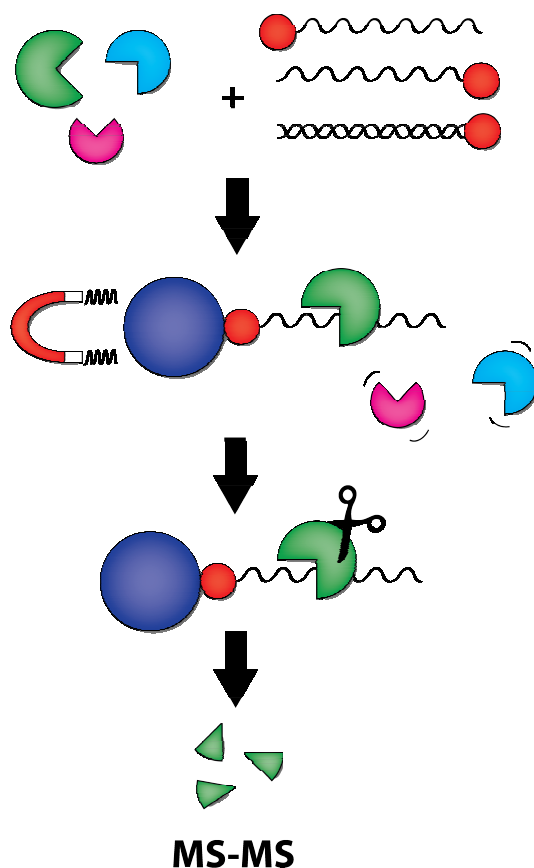


Fig. 6 – The solid-phase assay (SPA) used in **Paper II**. Proteins from the cellular lysate are represented on the top left as $\frac{3}{4}$ -circles, red circles represent biotin, blue circle represent paramagnetic streptavidin-coated beads, the scissor represent tryptic digestion prior to MS-MS.

In order to better imitate the *in vivo* DUS-binding conditions, the solid-phase assay (SPA) was designed (Fig. 6) (**Paper II**). Instead of detecting DNA- and DUS-specific binding on the immobilized proteins, the protein-DNA interaction was allowed to happen in liquid, allowing the native proteins to interact prior to the immobilization followed by identification and quantification by MS-MS. This method, utilizing DNA-covered magnetic-bead immobilization in solution with pull-down of potential targets and subsequent MS-MS identification and quantification using MaxQuant of the potential DUS-specific binding protein is to our knowledge the first of its kind. A number of different conditions were initially set up to optimize immobilization of the DNA- and DUS-specific binding protein(s). These alterations

included optimization of the DNA probe sequences and buffer conditions, improved targeting with the MS data. Using SDS-PAGE gels followed by protein staining, the background binding to the probes-less negative control, and the binding stringency for the DNA- and DUS-probes could be adjusted. Due to time-constraint the corresponding 5' dsDNA probes are lacking from the study. A correlation between the label-free quantity (LFQ) intensity values using the MaxQuant software and the protein abundance in the sample has previously been confirmed (de Godoy *et al.*, 2008; Lubner *et al.*, 2010), and it was assumed that the relative LFQ intensity represented the protein affinity to the respective probes. The final SPA and consecutive MS-MS analysis revealed that the TopA from Mc is the major DNA binding protein detected in this assay and may function as a mediator of the DUS specificity. Future *in vivo* studies are needed to confirm this putative TopA-activity. The methodological constraints by the MaxQuant threshold adopted from Målen *et al.* (Målen *et al.*, 2011) were tested by analysing the whole data-set using less restrictive criteria (lower relative intensity ratio and no lower peptide limit). With these less stringent thresholds the prominent DNA-binding affinity of TopA was confirmed, but more noise caused by proteins immobilized due to their extensive abundance was included. As such, the parameters and threshold used in this study were found to produce the most likely DNA and DUS-specific binding proteins. One question is why e.g. the abundantly expressed SSB was not identified with MS-MS even with the random ssDNA probe in the SPA assay.

In Mc, *topA* null-mutants are non-viable, *topA* thus being an essential gene (Ambur O.H., personal communication), in contrast to *H. influenzae* where the *topA* null-mutants are non-competent, but viable (Chandler & Smith, 1996). Studies in *E. coli* have revealed that loss of TopA is compensated by GyrA and GyrB (Sternglanz *et al.*, 1981; DiNardo *et al.*, 1982; Pruss *et al.*, 1982). If a homologous compensatory effect is found in Mc, this may offer new opportunities for indirect manipulations of TopA through GyrA and GyrB. Another methodological possibility is to create TopA chimeras, consisting of the functional domain from closely related *Neisseriaceae* with the putative preferences for other DUS dialects (Frye *et al.*, manuscript submitted). Or, if the recognition protein TopA is homologous to the uptake signal sequence (USS) -specific protein in *H. influenzae*, manipulations of the target gene may allow for changes in specificity from DUS to USS (or *vice versa*). Specific manipulations of the DNA binding-site of the TopA protein may also lead to non-specific binding, and

could confirm the sequence specificity of this potential DUS recognition protein. Over-expression of TopA and potential for his-tagging was not included in this study due to time-constraints. Band-shift analysis with DNA substrates with and without DUS revealed DNA-binding of TopA, but not DUS-specificity. The study included TopA antibody-specific Western blots of the SPA proteins, confirming the immobilization of TopA from both cellular fractions and isolated TopA. However, quantification of the proteins on the Western blots was not as sensitive as the MaxQuant analysis of the MS-MS, and did not reveal DUS-specificity over the random DNA probes, but did reveal strong abundance to the DNA-containing probes compared to the probe-less control. Nicking of ssDNA by recombinant meningococcal TopA has revealed that nicking occurred inside the DUS, but also at sites unrelated to DUS. The potential DUS-specificity of TopA therefore requires further dedicated studies to become conclusive.

4.5.2. Genome instability exhibited by PV

The PV assay proved to be useful for measuring instability in Mc, either as a proxy of the mutator activity of a BER mutant (**Paper III**) or as a measure of the instability of a repeat tract and influence of functional MMR (**Paper IV**). The study of the DNA glycosylase Fpg also focused on aspects of the BER functionality of the protein, including the affinity for DNA substrates representing different mutations (**Paper III**). Various substrates were used to test the affinity and binding of the Fpg protein to known BER substrates, such as faPy and 8oxoG opposite A, T, C and G. The results yielded findings similar to those in *E. coli*, indeed confirming that Mc Fpg was a prototypical DNA glycosylase of its kind. The PV assay was found to be sufficient to confirm the previous estimates of the intermediate mutator activity of single *fpg* null-mutant.

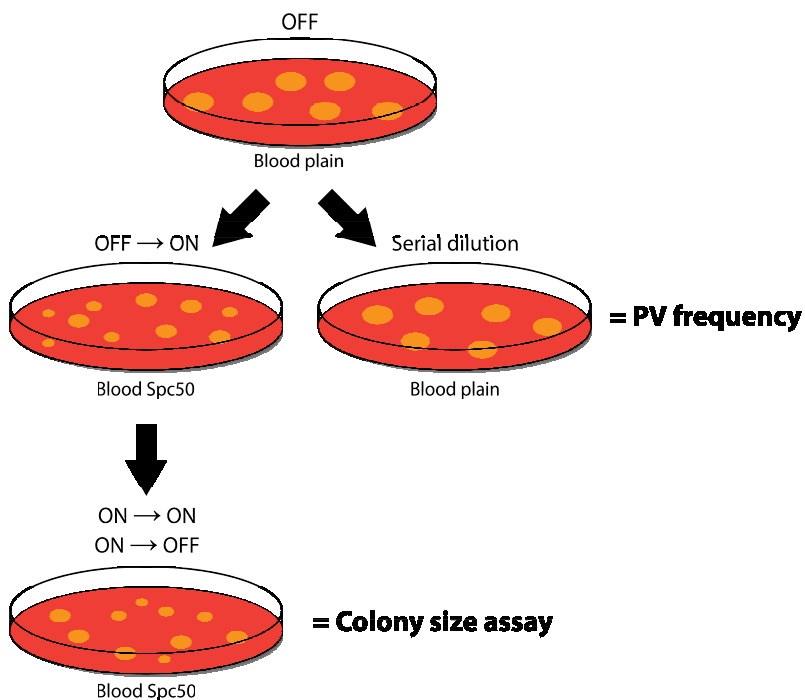


Fig. 7 – The phase variation (PV) method used in **Paper III and IV**. OFF and ON represent the different phases of the expression of the Spc resistance gene. PV frequency was calculated by the number of Spc^r colonies compared to the total number of colonies measured by serial dilutions on plain plates. Replated Spc^r colonies were used for analysis of the CFU colony size.

For the study of the different homopolymeric (polyG) tract lengths, in the wt and MMR background, a large panel of constructs for the PV assay was required (**Paper IV**). A substantial effort was invested into the construction of the panel of inherently non-stable polyG tracts, combined with the MMR null mutants, to a total of 18 constructs. Various approaches were proposed during the development of this study, which initially was aimed at determining the mutator effect of various DNA repair proteins (see **Paper III**). The idea of testing the effect of varying polyG tract lengths was conceived as this had never been studied in Mc before. Previous PV assays has shown that MMR proteins strongly influenced the PV frequency compared to other DNA repair pathway proteins, and therefore an MMR mutant was included in the assay to test how this would influence the PV rates of the various repeat tract lengths. Other PV studies have used mathematical models to calculate the instability of the phenotypic changes, the chances of back-switching, however, in this study we

utilized an *in vivo* approach to monitor ON \rightarrow OFF switching (Fig. 7). When performing the PV assay, it was observed that the CFUs were markedly smaller for the antibiotic-resistant phenotypes with longer repeat tracts than for the shorter. Quantitative measures of the CFU sizes confirmed this observation, and it was assumed that this size discrepancy was due to the increased instability of longer repeat tracts, thus giving rise to more extensive back-switching. This method is restricted in its resolution of the time of relatively late back-switching; however, it was inferred from comparing the PV frequencies and the CFU sizes that the back-switching occurred continuously. As such, it was not possible to quantify the amount of early back-switching colonies as these did not survive in either the PV frequency assay or the CFU size assay. Although this *in vivo* approach cannot quantify the early back-switchers, and could possibly be affected by other growth factors confounding the results, it was assumed that the difference in CFU size indeed was a measure of the instability of the repeat tracts responsible for the phenotypic expression of antibiotic resistance. New control strains harbouring non-homopolymeric and hence non-switching tracts encoding identical amino acids as the reporter strains are currently in preparation to control that the observed reduced survival of strains with long tracts are not a dominant negative effect of additional glycines.

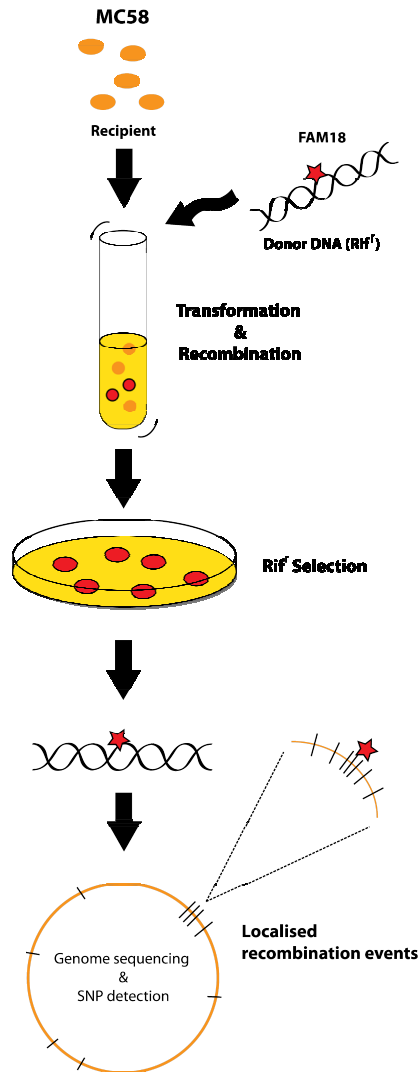


Fig. 8 – The method for sequencing of *Mc* following a transformation. The red star denotes the Rif^r mutation.

4.5.3. Recombination events delineated by genome sequencing

Whole genome sequencing has become more affordable and readily available for more applications of the methods. Thus, the whole genome sequencing of transformed *Mc* (**Paper V**) represents a study that could not have been done within reasonable time-frames and budgets just a few years ago.

A considerable effort was assigned to define the mutation and transformation rate of the library of *Mc* strain with available, sequenced genomes. Additionally, a

panel of various markers for transformation was tested, large inserts (antibiotic resistance cassettes, erythromycin and kanamycin resistance) and two antibiotic resistances acquired by spontaneous mutations of a single nucleotide (rifampin and streptomycin / spectinomycin resistance). Antibiotic resistance caused by a single nucleotide change rather than a large cassette (typically ~1 kb), was conceived as less invasive and hence preferred in this study. Among the two types of antibiotic resistance acquired by spontaneous mutations, Rif^r is currently most thoroughly investigated in *Mc*, and was therefore used. Transformation frequency was found to vary considerably between the four strains tested (three *Mc* strains and one *Gc* strain), however, the transformation frequency using the various markers were indifferent in each strain tested. Mutation frequency, measured as the ability to acquire the antibiotic resistances by spontaneous mutations, differed markedly between the strains tested, and allowed for an experimental design to decrease the chance of false positives by mutations in the recipient. In conclusion, the strain MC58 was selected as recipient for its high transformation frequency and low chance of spontaneous acquisition of the Rif^r. To be able to detect transformation in this experimental design, the DNA of the donor should be sufficiently genetically different from the recipient, MC58, but also sufficiently genetically similar to ensure homologous recombination. The *Mc* strain FAM18 was found to be suitably genetically similar / different to the recipient strain MC58 and readily acquired the spontaneous antibiotic resistance Rif^r and was used as the DNA donor (Fig. 8). Due to the promiscuity of *Mc*, foreign DNA from neisserial cells is readily taken up by transformation and incorporated through homologous recombination. To secure data from a range of transformation events, genomic DNA from 200 colonies were pooled and prepared for whole genome sequencing. For the bioinformatics analyses, the sequences were scrutinized for quality and quantity using reference-based assembly with Bowtie and Velvet, and restrictive parameters in Samtools to obtain the acquired SNPs in the transformed MC58 and the non-transformed controls. The SNP distribution surrounding the antibiotic resistance marker Rif^r was visualized as cumulative loss in expected SNPs against observed acquired SNPs, overlaid with a rolling mean of the genetic similarity between the donor and recipient strain and the positions of transcriptional terminators, yielding putative COHs and recombination fragment lengths.

5. Future prospects

The search for a DUS-specific DNA binding protein has proven challenging and yet unfulfilled, but has revealed a number of interesting findings along its way, such as the identification of significant DNA-specific binding proteins (**Papers I and II**), enabling novel findings on DNA metabolism in general. The search using *in vitro* experiments now seems exhausted, and future studies should resort to functional *in vivo* experiments using candidate DUS-specific binding proteins, such as those identified in **Paper II**. Identification of an active site and possible confirmation of DUS-specificity could be pursued using chimeric or complementation constructs from closely related species. Bioinformatics studies may also reveal novel possible DUS-specific binding candidates with the discovery of different DUS dialects within the *Neisseriaceae*.

Many of the known mutator genes in *Mc* has been thoroughly studied, especially the various components of the DNA repair pathways. **Paper III** confirms the proposed DNA repair activity of the *Mc* BER component, Fpg. *Mc* makes an ideal organism for mutation experiments, as the DNA repair pathways are well characterized, and as a model for DNA damage repair in other bacteria or eukaryotes. This bacterial-specific enzyme may also represent a new drug target.

The phase variation experiment (**Paper IV**) can be extended in multiple directions; other types of repeats can be compared (e.g. dinucleotide and heteropolymeric repeats), other phase variable loci and antibiotic markers utilized, comparison of other repair pathways and *Mc* strains. Another alternative direction could be to observe phase variation in *Mc in vivo*, treating infected mice with selected antibiotics in sub-optimal concentrations and observe the viability of the bacteria and the mice, and subsequent isolation and identification of the changes in the bacteria. Whole genome sequencing (see **Paper V**) can be applied to the phase variation experiment to observe the reach and genome-wide effect of the phase variation. Ultimately the model for tract instability can be applied to human diseases affected by nucleotide expansion and retraction, such as Huntington and ataxia.

Whole genome sequencing in general offers a new tool to assess genetic effects on the genome scale. Following the experiment already pursued (**Paper V**), other combinations of host- and donor-strain, of types of donor DNA can elaborate the story of neisserial transformation even further. The method could also be applied

to other bacterial species, to evaluate the genome-wide difference between various modes of HGT and even to monitor experimental evolution

There is growing interest in the study of genetic predisposition for disease, particularly cancer, and our model may inspire the application of similar experimental set-ups for the studies of mutation biases of other repair pathways. Insight into the genome dynamics of the pathogenic Mc may offer some new approaches to vaccinology and intervention of infections caused by a bacterium that is still responsible for a multitude of deaths worldwide. The translational value of the biological and evolutionary perspectives of the Mc life history to other organisms are obvious; as some aspects might be conserved in all domains of life, including humans, while some of the observed mechanisms that probably are bacterial (or neisserial) specific can be elaborated for diagnostic, preventive and therapeutic reasons.

6. References

- Alexander, H. L., Rasmussen, A. W. & Stojiljkovic, I. (2004). Identification of *Neisseria meningitidis* genetic loci involved in the modulation of phase variation frequencies. *Infection and Immunity* **72**, 6743-6747.
- Altmann, C. R., Solow-Cordero, D. E. & Chamberlin, M. J. (1994). RNA cleavage and chain elongation by *Escherichia coli* DNA-dependent RNA polymerase in a binary enzyme. RNA complex. *Proceedings of the National Academy of Sciences* **91**, 3784-3788.
- Ambur, O. H., Frye, S. A. & Tønjum, T. (2007). New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators. *The Journal of Bacteriology* **189**, 2077-2085.
- Ambur, O. H., Davidsen, T., Frye, S. A., Balasingham, S. V., Lagesen, K., Rognes, T. & Tønjum, T. (2009). Genome dynamics in major bacterial pathogens. *FEMS Microbiology Reviews* **33**, 453-470.
- Andrew, S. E., Paul Goldberg, Y., Kremer, B. & other authors (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nature Genetics* **4**, 398-403.
- Arnold, R., Galloway, Y., McNicholas, A. & O'Hallahan, J. (2011). Effectiveness of a vaccination programme for an epidemic of meningococcal B in New Zealand. *Vaccine* **29**, 7100-7106.
- Assalkhou, R., Balasingham, S., Collins, R. F., Frye, S. A., Davidsen, T., Benam, A. V., Bjørås, M., Derrick, J. P. & Tønjum, T. (2007). The outer membrane secretin PilQ from *Neisseria meningitidis* binds DNA. *Microbiology* **153**, 1593-1603.
- Balasingham, S. V., Collins, R. F., Assalkhou, R., Homberset, H., Frye, S. A., Derrick, J. P. & Tønjum, T. (2007). Interactions between the lipoprotein PilP and the secretin PilQ in *Neisseria meningitidis*. *The Journal of Bacteriology* **189**, 5716-5727.
- Banerjee, S., Chalissery, J., Bandey, I. & Sen, R. (2006). Rho-dependent transcription termination: more questions than answers. *The Journal of Microbiology* **44**, 11-22.
- Bayliss, C. D., Field, D. & Moxon, E. R. (2001). The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *The Journal of Clinical Investigation* **107**, 657-666.
- Bayliss, C. D., Dixon, K. M. & Moxon, E. R. (2004). Simple sequence repeats (microsatellites): mutational mechanisms and contributions to bacterial pathogenesis. A meeting review. *FEMS Immunology & Medical Microbiology* **40**, 11-19.
- Benam, A. V., Lång, E., Alfsnes, K., Fleckenstein, B., Rowe, A. D., Hovland, E., Ambur, O. H., Frye, S. A. & Tønjum, T. (2011). Structure-function relationships of the competence lipoprotein ComL and SSB in meningococcal transformation. *Microbiology* **157**, 1329-1342.
- Bennett, J., Jolley, K., Sparling, P. F., Saunders, N., Hart, C. A., Feavers, I. & Maiden, M. (2007). Species status of *Neisseria gonorrhoeae*: evolutionary and epidemiological inferences from multilocus sequence typing. *BMC Biology* **5**, 35.
- Bennett, J., Bentley, S., Vernikos, G., Quail, M., Cherevach, I., White, B., Parkhill, J. & Maiden, M. (2010). Independent evolution of the core and accessory gene sets in the genus *Neisseria*: insights gained from the genome of *Neisseria lactamica* isolate 020-06. *BMC Genomics* **11**, 652.
- Bentley, S. D., Vernikos, G. S., Snyder, L. A. S. & other authors (2007). Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genetics* **3**, e23.

- Bergé, M., Mortier-Barrière, I., Martin, B. & Claverys, J.-P. (2003).** Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Molecular Microbiology* **50**, 527-536.
- Bilukha, O. O. & Rosenstein, N. (2005).** Prevention and control of meningococcal disease. *MMWR Recomm Rep* **54**, 1-21.
- Biswas, G. D., Sox, T., Blackman, E. & Sparling, P. F. (1977).** Factors affecting genetic transformation of *Neisseria gonorrhoeae*. *The Journal of Bacteriology* **129**, 983-992.
- Bjune, G., Høiby, E. A., Grønnesby, J. K. & other authors (1991).** Effect of outer membrane vesicle vaccine against group B meningococcal disease in Norway. *The Lancet* **338**, 1093-1096.
- Bohannon, B. J. M. & Lenski, R. E. (2000).** Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecology Letters* **3**, 362-377.
- Boisier, P., Djibo, S., Sidikou, F., Mindadou, H., Kairo, K. K., Djibo, A., Goumbi, K. & Chanteau, S. (2005).** Epidemiological patterns of meningococcal meningitis in Niger in 2003 and 2004: under the threat of *N. meningitidis* serogroup W135. *Tropical Medicine & International Health* **10**, 435-443.
- Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. (2005).** Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551-2561.
- Borud, B., Aas, F. E., Vik, Å., Winther-Larsen, H. C., Egge-Jacobsen, W. & Koomey, M. (2010).** Genetic, structural, and antigenic analyses of glycan diversity in the O-linked protein glycosylation systems of human *Neisseria* species. *The Journal of Bacteriology* **192**, 2816-2829.
- Brandtzaeg, P., Bjerre, A., Øvstebø, R., Brusletto, B., Jøø, G. B. & Kierulf, P. (2001).** Invited review: *Neisseria meningitidis* lipopolysaccharides in human pathology. *Journal of Endotoxin Research* **7**, 401-420.
- Brandtzaeg, P. & van Deuren, M. (2002).** Current concepts in the role of the host response in *Neisseria meningitidis* septic shock. *Current Opinion in Infectious Diseases* **15**, 247-252.
- Budroni, S., Siena, E., Hotopp, J. C. D. & other authors (2011).** *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proceedings of the National Academy of Sciences* **108**, 4494-4499.
- Buisine, N., Tang, C. M. & Chalmers, R. (2002).** Transposon-like Correia elements: structure, distribution and genetic exchange between pathogenic *Neisseria* sp. *FEBS Letters* **522**, 52-58.
- Campbell, E. A., Choi, S. Y. & Masure, H. R. (1998).** A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis. *Molecular Microbiology* **27**, 929-939.
- Cassio de Moraes, J., Camargo, M. C. C., Rossetto Hidalgo, N. T. & other authors (1992).** Protective efficacy of a serogroup B meningococcal vaccine in Sao Paulo, Brazil. *The Lancet* **340**, 1074-1078.
- Caugant, D. A., Zollinger, W. D., Mocca, L. F., Frasch, C. E., Whittam, T. S., Frøholm, L. O. & Selander, R. K. (1987).** Genetic relationships and clonal population structure of serotype 2 strains of *Neisseria meningitidis*. *Infection and Immunity* **55**, 1503-1512.
- Caugant, D. A. (1998).** Population genetics and molecular epidemiology of *Neisseria meningitidis*. *APMIS* **106**, 505-525.
- Caugant, D. A. (2008).** Genetics and evolution of *Neisseria meningitidis*: Importance for the epidemiology of meningococcal disease. *Infection, Genetics and Evolution* **8**, 558-565.

- Chamot-Rooke, J., Mikaty, G., Malosse, C. & other authors (2011).** Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination. *Science* **331**, 778-782.
- Champoux, J. J. (2001).** DNA topoisomerases: structure, function, and mechanism. *Annual Review of Biochemistry* **70**, 369-413.
- Chandler, M. S. & Smith, R. A. (1996).** Characterization of the *Haemophilus influenzae* topA locus: DNA topoisomerase I is required for genetic competence. *Gene* **169**, 25-31.
- Chen, C.-C., Hsia, K.-C., Huang, C.-T., Wong, W.-W., Yen, M.-Y., Li, L.-H., Lin, K.-Y., Chen, K.-W. & Li, S.-Y. (2011).** Draft genome sequence of a dominant, multidrug-resistant *Neisseria gonorrhoeae* strain, TCDC-NG08107, from a sexual group at high risk of acquiring human immunodeficiency virus infection and Syphilis. *The Journal of Bacteriology* **193**, 1788-1789.
- Chen, I. & Gotschlich, E. C. (2001).** ComE, a competence protein from *Neisseria gonorrhoeae* with DNA-binding activity. *The Journal of Bacteriology* **183**, 3160-3168.
- Chen, I. & Dubnau, D. (2003).** DNA transport during transformation. *Front Biosci* **1**, 544-556.
- Chen, I. & Dubnau, D. (2004).** DNA uptake during bacterial transformation. *Nat Rev Micro* **2**, 241-249.
- Chen, I., Christie, P. J. & Dubnau, D. (2005).** The ins and outs of DNA transfer in bacteria. *Science* **310**, 1456-1460.
- Chung, G. T., Yoo, J. S., Oh, H. B., Lee, Y. S., Cha, S. H., Kim, S. J. & Yoo, C. K. (2008).** Complete genome sequence of *Neisseria gonorrhoeae* NCCP11945. *The Journal of Bacteriology* **190**, 6035-6036.
- Claverys, J.-P. & Martin, B. (2003).** Bacterial 'competence' genes: signatures of active transformation, or only remnants? *Trends in Microbiology* **11**, 161-165.
- Claverys, J.-P., Martin, B. & Polard, P. (2009).** The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiology Reviews* **33**, 643-656.
- Collins, R. F., Frye, S. A., Balasingham, S., Ford, R. C., Tønjum, T. & Derrick, J. P. (2005).** Interaction with Type IV pili induces structural changes in the bacterial outer membrane secretin PilQ. *Journal of Biological Chemistry* **280**, 18923-18930.
- Danner, D. B., Deich, R. A., Sisco, K. L. & Smith, H. O. (1980).** An eleven-base-pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. *Gene* **11**, 311-318.
- Davidson, T., Rødland, E. A., Lagesen, K., Seeberg, E., Rognes, T. & Tønjum, T. (2004).** Biased distribution of DNA uptake sequences towards genome maintenance genes. *Nucleic Acids Research* **32**, 1050-1058.
- Davidson, T., Bjørås, M., Seeberg, E. C. & Tønjum, T. (2005).** Antimutator role of DNA glycosylase MutY in pathogenic *Neisseria* species. *Journal of Bacteriology* **187**, 2801-2809.
- Davidson, T. & Tønjum, T. (2006).** Meningococcal genome dynamics. *Nat Rev Micro* **4**, 11-22.
- Davidson, T., Tuven, H. K., Bjoras, M., Rodland, E. A. & Tonjum, T. (2007).** Genetic interactions of DNA repair pathways in the pathogen *Neisseria meningitidis*. *The Journal of Bacteriology* **189**, 5728-5737.
- Dawkins, R. & Krebs, J. R. (1979).** Arms races between and within species. *Proceedings of the Royal Society of London Series B Biological Sciences* **205**, 489-511.
- De Bolle, X., Bayliss, C. D., Field, D., Van De Ven, T., Saunders, N. J., Hood, D. W. & Moxon, E. R. (2000).** The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase

variation rate of a gene with homology to type III DNA methyltransferases. *Molecular Microbiology* **35**, 211-222.

de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Frohlich, F., Walther, T. C. & Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251-1254.

Dehio, C., Gray-Owen, S. D. & Meyer, T. F. (1998). The role of neisserial Opa proteins in interactions with host cells. *Trends in Microbiology* **6**, 489-495.

Demple, B. & Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annual Review of Biochemistry* **63**, 915-948.

DeVoe, I. W. (1982). The meningococcus and mechanisms of pathogenicity. *Microbiology Review* **46**, 162-190.

Didelot, X. & Falush, D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**, 1251-1266.

DiNardo, S., Voelkel, K. A., Sternglanz, R., Reynolds, A. E. & Wright, A. (1982). *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. *Cell* **31**, 43-51.

Donnelly, J., Medini, D., Boccadifuoco, G. & other authors (2010). Qualitative and quantitative assessment of meningococcal antigens to evaluate the potential strain coverage of protein-based vaccines. *Proceedings of the National Academy of Sciences* **107**, 19490-19495.

Drake, J. W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proceedings of the National Academy of Sciences* **88**, 7160-7164.

Dubnau, D. (1999). DNA uptake in bacteria. *Annual Review of Microbiology* **53**, 217-244.

Duffin, P. M. & Seifert, H. S. (2010). DNA uptake sequence-mediated enhancement of transformation in *Neisseria gonorrhoeae* is strain dependent. *The Journal of Bacteriology* **192**, 4436-4444.

Elkins, C., Thomas, C. E., Seifert, H. S. & Sparling, P. F. (1991). Species-specific uptake of DNA by gonococci is mediated by a 10-base-pair sequence. *The Journal of Bacteriology* **173**, 3911-3913.

Evans, C. M., Pratt, C. B., Matheson, M., Vaughan, T. E., Findlow, J., Borrow, R., Gorrings, A. R. & Read, R. C. (2011). Nasopharyngeal colonization by *Neisseria lactamica* and induction of protective immunity against *Neisseria meningitidis*. *Clinical Infectious Diseases* **52**, 70-77.

Facius, D. & Meyer, T. F. (1993). Novel determinant (comA) essential for natural transformation competence in *Neisseria gonorrhoeae* and the effect of a comA defect on pilin variation. *Molecular Microbiology* **10**, 699-712.

Farnham, P. J. & Platt, T. (1981). Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. *Nucleic Acids Research* **9**, 563-577.

Feil, E. J., Maiden, M. C., Achtman, M. & Spratt, B. G. (1999). The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. *Molecular Biology and Evolution* **16**, 1496-1502.

Feil, E. J., Enright, M. C. & Spratt, B. G. (2000). Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between *Neisseria meningitidis* and *Streptococcus pneumoniae*. *Research in Microbiology* **151**, 465-469.

Feil, E. J. & Spratt, B. G. (2001). Recombination and the population structures of bacterial pathogens. *Annual Review of Microbiology* **55**, 561-590.

- Finne, J., Leinonen, M. & Mäkelä, P. H. (1983).** Antigenic similarities between brain components and bacteria causing meningitis: implications for vaccine development and pathogenesis. *The Lancet* **322**, 355-357.
- Finney, M., Vaughan, T., Taylor, S. & other authors (2008).** Characterization of the key antigenic components and pre-clinical immune responses to a meningococcal disease vaccine based on *Neisseria lactamica* outer membrane vesicles. *Human Vaccines* **4**, 23-30.
- Forgor, A. A., Leimkugel, J., Hodgson, A., Bugri, A., Dangy, J.-P., Gagneux, S., Smith, T. & Pluschke, G. (2005).** Emergence of W135 meningococcal meningitis in Ghana. *Tropical Medicine & International Health* **10**, 1229-1234.
- Freudenreich, C. H., Kantrow, S. M. & Zakian, V. A. (1998).** Expansion and length-dependent fragility of CTG repeats in yeast. *Science* **279**, 853-856.
- Frye, S. A., Assalkhou, R., Collins, R. F., Ford, R. C., Petersson, C., Derrick, J. P. & Tonjum, T. (2006).** Topology of the outer-membrane secretin PilQ from *Neisseria meningitidis*. *Microbiology* **152**, 3751-3764.
- Fussenegger, M., Facius, D., Meier, J. & Meyer, T. F. (1996a).** A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of *Neisseria gonorrhoeae*. *Molecular Microbiology* **19**, 1095-1105.
- Fussenegger, M., Kahrs, A. F., Facius, D. & Meyer, T. F. (1996b).** Tetrapac (tpc), a novel genotype of *Neisseria gonorrhoeae* affecting epithelial cell invasion, natural transformation competence and cell separation. *Molecular Microbiology* **19**, 1357-1372.
- Fussenegger, M., Rudel, T., Barten, R., Ryll, R. & Meyer, T. F. (1997).** Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae* - a review. *Gene* **192**, 125-134.
- Girard, M. P., Preziosi, M.-P., Aguado, M.-T. & Kieny, M. P. (2006).** A review of vaccine research and development: meningococcal disease. *Vaccine* **24**, 4692-4700.
- Giuliani, M. M., Adu-Bobie, J., Comanducci, M. & other authors (2006).** A universal vaccine for serogroup B meningococcus. *Proceedings of the National Academy of Sciences* **103**, 10834-10839.
- Goodman, S. D. & Scocca, J. J. (1988).** Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences* **85**, 6982-6986.
- Gotschlich, E. C., Goldschneider, I. & Artenstein, M. S. (1969).** Human immunity to the meningococcus. *The Journal of Experimental Medicine* **129**, 1385-1395.
- Graves, J. F., Biswas, G. D. & Sparling, P. F. (1982).** Sequence-specific DNA uptake in transformation of *Neisseria gonorrhoeae*. *Journal of Bacteriology* **152**, 1071-1077.
- Hammerschmidt, S., Müller, A., Sillmann, H. & other authors (1996).** Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (siaD): correlation with bacterial invasion and the outbreak of meningococcal disease. *Molecular Microbiology* **20**, 1211-1220.
- Hardy, S. J., Christodoulides, M., Weller, R. O. & Heckels, J. E. (2000).** Interactions of *Neisseria meningitidis* with cells of the human meninges. *Molecular Microbiology* **36**, 817-829.
- Henderson, I. R., Owen, P. & Nataro, J. P. (1999).** Molecular switches — the ON and OFF of bacterial phase variation. *Molecular Microbiology* **33**, 919-932.
- Håvarstein, L. S., Coomaraswamy, G. & Morrison, D. A. (1995).** An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proceedings of the National Academy of Sciences* **92**, 11140-11144.

Iyer, R. R., Pluciennik, A., Burdett, V. & Modrich, P. L. (2006). DNA mismatch repair: functions and mechanisms. *ChemInform* **37**, no.

Jarva, H., Ram, S., Vogel, U., Blom, A. M. & Meri, S. (2005). Binding of the complement inhibitor C4bp to serogroup B *Neisseria meningitidis*. *The Journal of Immunology* **174**, 6299-6307.

Johnsborg, O., Eldholm, V. & Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Research in Microbiology* **158**, 767-778.

Johnson, T. L., Abendroth, J., Hol, W. G. J. & Sandkvist, M. (2006). Type II secretion: from structure to function. *FEMS Microbiology Letters* **255**, 175-186.

Jolley, K., Chan, M.-S. & Maiden, M. (2004). mlstdbNet - distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics* **5**, 86.

Joseph, B., Schneiker-Bekel, S., Schramm-Glück, A. & other authors (2010). Comparative genome biology of a serogroup B carriage and disease strain supports a polygenic nature of meningococcal virulence. *The Journal of Bacteriology* **192**, 5363-5377.

Jysum, K. & Lie, S. (1965). Genetic factors determining competence in transformation of *Neisseria meningitidis*. 1. A permanent loss of competence. *Acta Pathol Microbiol Scand* **63**, 306-316.

Jäkel, A., Plested, J. S., Hoe, J. C. & other authors (2008). Naturally-occurring human serum antibodies to inner core lipopolysaccharide epitopes of *Neisseria meningitidis* protect against invasive meningococcal disease caused by isolates displaying homologous inner core structures. *Vaccine* **26**, 6655-6663.

Kadyrov, F. A., Dzantiev, L., Constantin, N. & Modrich, P. (2006). Endonucleolytic function of MutLa in human mismatch repair. *Cell* **126**, 297-308.

Kiefer, F. (1896). Zur differential diagnose des erregers der epidemischen cerebrospinalmeningitis und der gonorrhoea. *Berl Klin Wochenschr* **33**, 628-630.

Kingsford, C., Ayanbule, K. & Salzberg, S. (2007). Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biology* **8**, R22.

Klee, S. R., Nassif, X., Kusecek, B., Merker, P., Beretti, J.-L., Achtman, M. & Tinsley, C. R. (2000). Molecular and biological analysis of eight genetic islands that distinguish *Neisseria meningitidis* from the closely related pathogen *Neisseria gonorrhoeae*. *Infection and Immunity* **68**, 2082-2095.

Koomey, J. M. & Falkow, S. (1987). Cloning of the *recA* gene of *Neisseria gonorrhoeae* and construction of gonococcal *recA* mutants. *The Journal of Bacteriology* **169**, 790-795.

Koomey, J. M. (1998). Competence for natural transformation in *Neisseria gonorrhoeae*: a model system for studies of horizontal gene transfer. *APMIS Suppl* **84**, 56-61.

Koomey, M., Gotschlich, E. C., Robbins, K., Bergstrom, S. & Swanson, J. (1987). Effects of *recA* mutations on pilus antigenic variation and phase transitions in *Neisseria gonorrhoeae*. *Genetics* **117**, 391-398.

Kristiansen, P. A., Diomandé, F., Wei, S. C. & other authors (2011). Baseline meningococcal carriage in Burkina Faso before the introduction of a meningococcal serogroup A conjugate vaccine. *Clinical and Vaccine Immunology* **18**, 435-443.

Lawrence, J. G., Hendrix, R. W. & Casjens, S. (2001). Where are the pseudogenes in bacterial genomes? *Trends in Microbiology* **9**, 535-540.

- LeCuyer, B. E., Criss, A. K. & Seifert, H. S. (2010).** Genetic characterization of the nucleotide excision repair system of *Neisseria gonorrhoeae*. *Journal of Bacteriology* **192**, 665-673.
- Li, L. & Lu, A.-L. (2003).** The C - terminal domain of *Escherichia coli* MutY is involved in DNA binding and glycosylase activities. *Nucleic Acids Research* **31**, 3038-3049.
- Li, X., Wright, P. M. & Lu, A. L. (2000).** The C-terminal domain of MutY glycosylase determines the 7,8-dihydro-8-oxo-guanine specificity and is crucial for mutation avoidance. *Journal of Biological Chemistry* **275**, 8448-8455.
- Lindahl, T. (1993).** Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715.
- Lindner, C., Nijland, R., van Hartskamp, M., Bron, S., Hamoen, L. W. & Kuipers, O. P. (2004).** Differential expression of two paralogous genes of *Bacillus subtilis* encoding single-stranded DNA binding protein. *The Journal of Bacteriology* **186**, 1097-1105.
- Little, C. T. S. & Vrijenhoek, R. C. (2003).** Are hydrothermal vent animals living fossils? *Trends in Ecology & Evolution* **18**, 582-588.
- Lorenz, M. G. & Wackernagel, W. (1994).** Bacterial gene transfer by natural genetic transformation in the environment. *Microbiology and Molecular Biology Reviews* **58**, 563-602.
- Luber, C. A., Cox, J., Lauterbach, H. & other authors (2010).** Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279-289.
- Lucidarme, J., Newbold, L. S., Findlow, J., Gilchrist, S., Gray, S. J., Carr, A. D., Hewitt, J., Kaczmarek, E. B. & Borrow, R. (2011).** Molecular targets in meningococci: efficient routine characterization and optimal outbreak investigation in conjunction with routine surveillance of the meningococcal group B vaccine candidate, fHBP. *Clinical and Vaccine Immunology* **18**, 194-202.
- Lång, E., Haugen, K., Fleckenstein, B., Homberset, H., Frye, S. A., Ambur, O. H. & Tønjum, T. (2009).** Identification of neisserial DNA binding components. *Microbiology* **155**, 852-862.
- Macfadyen, L. P., Chen, D., Vo, H. C., Liao, D., Sinotte, R. & Redfield, R. J. (2001).** Competence development by *Haemophilus influenzae* is regulated by the availability of nucleic acid precursors. *Molecular Microbiology* **40**, 700-707.
- Macia, M. D., Blanquer, D., Togores, B., Sauleda, J., Perez, J. L. & Oliver, A. (2005).** Hypermutation is a key factor in development of multiple-antimicrobial resistance in *Pseudomonas aeruginosa* strains causing chronic lung infections. *Antimicrobial Agents and Chemotherapy* **49**, 3382-3386.
- Maiden, M. C. J., Bygraves, J. A., Feil, E. & other authors (1998).** Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences* **95**, 3140-3145.
- Makarova, K. S., Haft, D. H., Barrangou, R. & other authors (2011).** Evolution and classification of the CRISPR-Cas systems. *Nat Rev Micro* **9**, 467-477.
- Marsh, J. W., Shutt, K. A., Pajon, R. & other authors (2011).** Diversity of factor H-binding protein in *Neisseria meningitidis* carriage isolates. *Vaccine* **29**, 6049-6058.
- Martin, P., Ven, T. v. d., Mouchel, N., Jeffries, A. C., Hood, D. W. & Moxon, E. R. (2003).** Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. *Molecular Microbiology* **50**, 245-257.
- Martin, P., Sun, L., Hood, D. W. & Moxon, E. R. (2004).** Involvement of genes of genome maintenance in the regulation of phase variation frequencies in *Neisseria meningitidis*. *Microbiology* **150**, 3001-3012.

- Massari, P., Ho, Y. & Wetzler, L. M. (2000).** *Neisseria meningitidis* porin PorB interacts with mitochondria and protects cells from apoptosis. *Proceedings of the National Academy of Sciences* **97**, 9070-9075.
- Mathis, L. S. & Scocca, J. J. (1982).** *Haemophilus influenzae* and *Neisseria gonorrhoeae* recognize different specificity determinants in the DNA uptake step of genetic transformation. *Journal of General Microbiology* **128**, 1159-1161.
- Mehr, I. J. & Seifert, H. S. (1998).** Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair. *Molecular Microbiology* **30**, 697-710.
- Modrich, P. & Lahue, R. (1996).** Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry* **65**, 101-133.
- Mojica, F. J. M., Díez-Villaseñor, C. s., García-Martínez, J. & Soria, E. (2005).** Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* **60**, 174-182.
- Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. (1994).** Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology* **4**, 24-33.
- Moxon, R., Bayliss, C. & Hood, D. (2006).** Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annual Review of Genetics* **40**, 307-333.
- Målen, H., De Souza, G., Pathak, S., Sjøfteland, T. & Wiker, H. (2011).** Comparison of membrane proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra strains. *BMC Microbiology* **11**, 18.
- Nagorska, K., Silhan, J., Li, Y., Pelicic, V., Freemont, P. S., Baldwin, G. S. & Tang, C. M. (2012).** A network of enzymes involved in repair of oxidative DNA damage in *Neisseria meningitidis*. *Molecular Microbiology* **83**, 1064-1079.
- Napolitano, R., Janel-Bintz, R., Wagner, J. & Fuchs, R. P. P. (2000).** All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J* **19**, 6259-6265.
- Nassif, X., Lowy, J., Stenberg, P., O'Gaora, P., Ganji, A. & So, M. (1993).** Antigenic variation of pilin regulates adhesion of *Neisseria meningitidis* to human epithelial cells. *Molecular Microbiology* **8**, 719-725.
- Nassif, X., Pujol, C., Morand, P. & Eugène, E. (1999).** Interactions of pathogenic *Neisseria* with host cells. Is it possible to assemble the puzzle? *Molecular Microbiology* **32**, 1124-1132.
- Nassif, X. (2000).** A furtive pathogen revealed. *Science* **287**, 1767-1768.
- Nouspikel, T. & Hanawalt, P. C. (2002).** DNA repair in terminally differentiated cells. *DNA Repair* **1**, 59-75.
- Nunn, D. (1999).** Bacterial Type II protein export and pilus biogenesis: more than just homologies? *Trends in Cell Biology* **9**, 402-408.
- Oliver, A., Cantón, R., Campo, P., Baquero, F. & Blázquez, J. (2000).** High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* **288**, 1251-1253.
- Oliver, K. J., Reddin, K. M., Bracegirdle, P., Hudson, M. J., Borrow, R., Feavers, I. M., Robinson, A., Cartwright, K. & Gorringe, A. R. (2002).** *Neisseria lactamica* protects against experimental meningococcal infection. *Infection and Immunity* **70**, 3621-3626.
- Orr, H. T., Chung, M.-y., Banfi, S. & other authors (1993).** Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nature Genetics* **4**, 221-226.

- Pajon, R., Beernink, P. T., Harrison, L. H. & Granoff, D. M. (2010).** Frequency of factor H-binding protein modular groups and susceptibility to cross-reactive bactericidal activity in invasive meningococcal isolates. *Vaccine* **28**, 2122-2129.
- Parkhill, J., Achtman, M., James, K. D. & other authors (2000).** Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* **404**, 502-506.
- Perrin, A., Bonacorsi, S., Carbonnelle, E., Talibi, D., Dessen, P., Nassif, X. & Tinsley, C. (2002).** Comparative genomics identifies the genetic islands that distinguish *Neisseria meningitidis*, the agent of cerebrospinal meningitis, from other *Neisseria* species. *Infection and Immunity* **70**, 7063-7072.
- Pillon, M. C., Lorenowicz, J. J., Uckelmann, M. & other authors (2010).** Structure of the endonuclease domain of MutL: unlicensed to cut. *Molecular Cell* **39**, 145-151.
- Pourcel, C., Salvignol, G. & Vergnaud, G. (2005).** CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653-663.
- Pruss, G. J., Manes, S. H. & Drlica, K. (1982).** *Escherichia coli* DNA topoisomerase I mutants: increased supercoiling is corrected by mutations near gyrase genes. *Cell* **31**, 35-42.
- Pugsley, A. P. (1993).** The complete general secretory pathway in gram-negative bacteria. *Microbiology and Molecular Biology Reviews* **57**, 50-108.
- Pujol, C. I., Eugène, E., Morand, P. & Nassif, X. (2000).** Do pathogenic *Neisseriae* need several ways to modify the host cell cytoskeleton? *Microbes and Infection* **2**, 821-827.
- Rappuoli, R. (2001).** Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine* **19**, 2688-2691.
- Richardson, A. R. & Stojiljkovic, I. (2001).** Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*. *Molecular Microbiology* **40**, 645-655.
- Richardson, A. R., Yu, Z., Popovic, T. & Stojiljkovic, I. (2002).** Mutator clones of *Neisseria meningitidis* in epidemic serogroup A disease. *Proceedings of the National Academy of Sciences* **99**, 6103-6107.
- Roberts, L. (2010).** The beginning of the end for Africa's devastating meningitis outbreaks? *Science* **330**, 1466-1467.
- Rosenberg, S. M. (2001).** Evolving responsively: adaptive mutation. *Nature Reviews Genetics* **2**, 504-515.
- Rudel, T., Schmid, A., Benz, R., Kolb, H.-A., Lang, F. & Meyer, T. F. (1996).** Modulation of *Neisseria* porin (PorB) by cytosolic ATP/GTP of target cells: parallels between pathogen accommodation and mitochondrial endosymbiosis. *Cell* **85**, 391-402.
- Rusniok, C., Vallenet, D., Floquet, S. & other authors (2009).** NeMeSys: a biological resource for narrowing the gap between sequence and function in the human pathogen *Neisseria meningitidis*. *Genome Biology* **10**, 1-13.
- Sadarangani, M., Pollard, A. J. & Gray-Owen, S. D. (2011).** Opa proteins and CEACAMs: pathways of immune engagement for pathogenic *Neisseria*. *FEMS Microbiology Reviews* **35**, 498-514.
- Sardiñas, G., Reddin, K., Pajon, R. & Gorringe, A. (2006).** Outer membrane vesicles of *Neisseria lactamica* as a potential mucosal adjuvant. *Vaccine* **24**, 206-214.
- Saunders, N. J., Jeffries, A. C., Peden, J. F., Hood, D. W., Tettelin, H., Rappuoli, R. & Moxon, E. R. (2000).** Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. *Molecular Microbiology* **37**, 207-215.

- Saunders, N. J., Moxon, E. R. & Gravenor, M. B. (2003).** Mutation rates: estimating phase variation rates when fitness differences are present and their impact on population structure. *Microbiology* **149**, 485-495.
- Schoen, C., Blom, J., Claus, H. & other authors (2008).** Whole-genome comparison of disease and carriage strains provides insights into virulence evolution in *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences* **105**, 3473-3478.
- Schoen, C., Weber-Lehmann, J., Blom, J., Joseph, B., Goesmann, A., Strittmatter, A. & Frosch, M. (2011).** Whole-genome sequence of the transformable *Neisseria meningitidis* serogroup A strain WUE2594. *The Journal of Bacteriology* **193**, 2064-2065.
- Schofield, M. J. & Hsieh, P. (2003).** DNA mismatch repair: molecular mechanisms and biological function. *Annual Review of Microbiology* **57**, 579-608.
- Seeberg, E., Eide, L. & Bjørås, M. (1995).** The base excision repair pathway. *Trends in Biochemical Sciences* **20**, 391-397.
- Seib, K. L., Serruto, D., Oriente, F., Delany, I., Adu-Bobie, J., Veggi, D., Aricò, B., Rappuoli, R. & Pizza, M. (2009).** Factor H-binding protein is important for meningococcal survival in human whole blood and serum and in the presence of the antimicrobial peptide LL-37. *Infection and Immunity* **77**, 292-299.
- Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. & Keck, J. L. (2008).** SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology* **43**, 289-318.
- Shinde, D., Lai, Y., Sun, F. & Arnheim, N. (2003).** Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)_n and (A/T)_n microsatellites. *Nucleic Acids Research* **31**, 974-980.
- Singh, R. N. & Pitale, M. P. (1968).** Competence and deoxyribonucleic acid uptake in *Bacillus subtilis*. *Journal of Bacteriology* **95**, 864-866.
- Sisco, K. L. & Smith, H. O. (1979).** Sequence-specific DNA uptake in *Haemophilus* transformation. *Proceedings of the National Academy of Sciences* **76**, 972-976.
- Snyder, L., Butcher, S. A. & Saunders, N. J. (2001).** Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic *Neisseria* spp. *Microbiology* **147**, 2321-2332.
- Snyder, L. A. S., Cole, J. & Pallen, M. (2009).** Comparative analysis of two *Neisseria gonorrhoeae* genome sequences reveals evidence of mobilization of *Correia* repeat enclosed elements and their role in regulation. *BMC Genomics* **10**, 70.
- Soltis, P. S., Soltis, D. E., Savolainen, V., Crane, P. R. & Barraclough, T. G. (2002).** Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. *Proceedings of the National Academy of Sciences* **99**, 4430-4435.
- Sow, S. O., Okoko, B. J., Diallo, A. & other authors (2011).** Immunogenicity and safety of a meningococcal A conjugate vaccine in Africans. *The New England Journal of Medicine* **364**, 2293-2304.
- Sparling, P. F. (1966).** Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. *The Journal of Bacteriology* **92**, 1364-1371.
- Spratt, B. G., Hanage, W. P. & Feil, E. J. (2001).** The relative contributions of recombination and point mutation to the diversification of bacterial clones. *Current Opinion in Microbiology* **4**, 602-606.
- Stephens, D. S. (2007).** Conquering the meningococcus. *FEMS Microbiology Reviews* **31**, 3-14.

Stephens, D. S., Greenwood, B. & Brandtzaeg, P. (2007). Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*. *The Lancet* **369**, 2196-2210.

Sternglanz, R., DiNardo, S., Voelkel, K. A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L. & Wang, J. C. (1981). Mutations in the gene coding for *Escherichia coli* DNA topoisomerase I affect transcription and transposition. *Proceedings of the National Academy of Sciences* **78**, 2747-2751.

Streisinger, G. & Owen, J. (1985). Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics* **109**, 633-659.

Swanson, J. (1973). Studies on gonococcus infection - IV. pili: Their role in attachment of gonococci to tissue culture cells. *The Journal of Experimental Medicine* **137**, 571-589.

Swanson, J. (1978). Studies on gonococcus infection. XII. Colony color and opacity variants of gonococci. *Infection and Immunity* **19**, 320-331.

Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. & Nishimura, S. (1991). 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proceedings of the National Academy of Sciences* **88**, 4690-4694.

Tettelin, H., Saunders, N. J., Heidelberg, J. & other authors (2000). Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science* **287**, 1809-1815.

Touzain, F., Petit, M.-A., Schbath, S. & Karoui, M. E. (2011). DNA motifs that sculpt the bacterial chromosome. *Nat Rev Micro* **9**, 15-26.

Tran, H. T., Keen, J. D., Krickler, M., Resnick, M. A. & Gordenin, D. A. (1997). Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Molecular and Cellular Biology* **17**, 2859-2865.

Treangen, T., Ambur, O., Tønjum, T. & Rocha, E. (2008). The impact of the neisserial DNA uptake sequences on genome evolution and stability. *Genome Biology* **9**, R60.

Tzeng, Y.-L. & Stephens, D. S. (2000). Epidemiology and pathogenesis of *Neisseria meningitidis*. *Microbes and Infection* **2**, 687-700.

Tønjum, T., Freitag, N. E., Namork, E. & Koomey, M. (1995). Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. *Molecular Microbiology* **16**, 451-464.

Tønjum, T., Caugant, D. A., Dunham, S. A. & Koomey, M. (1998). Structure and function of repetitive sequence elements associated with a highly polymorphic domain of the *Neisseria meningitidis* PilQ protein. *Molecular Microbiology* **29**, 111-124.

Tønjum, T. & Seeberg, E. (2001). Microbial fitness and genome dynamics. *Trends in Microbiology* **9**, 356-358.

van der Woude, M. W. & Bäumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clinical Microbiology Reviews* **17**, 581-611.

van Deuren, M., Brandtzaeg, P. & van der Meer, J. W. M. (2000). Update on meningococcal disease with emphasis on pathogenesis and clinical management. *Clinical Microbiology Reviews* **13**, 144-166.

van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory* **1**, 1-30.

Viesseux, G. (1806). Memoire sur la maladie qui regné à Geneve au printemps de 1805. In *J Méd Chir Pharmacol*, pp. 50-53. Bruxelles.

Vignon, G., Kohler, R., Larquet, E., Giroux, S., Prevost, M.-C., Roux, P. & Pugsley, A. P. (2003). Type IV-like pili formed by the Type II secretion: specificity, composition, bundling, polar localization, and surface presentation of peptides. *The Journal of Bacteriology* **185**, 3416-3428.

Vik, Å., Aas, F. E., Anonsen, J. H., Bilsborough, S., Schneider, A., Egge-Jacobsen, W. & Koomey, M. (2009). Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences* **106**, 4447-4452.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O. & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature* **416**, 389-395.

Watt, V. M., Ingles, C. J., Urdea, M. S. & Rutter, W. J. (1985). Homology requirements for recombination in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **82**, 4768-4772.

Weichselbaum, A. (1887). Ueber die Aetiologie der akuten Meningitis Cerebrospinalis. *Fortschr Med* **5**, 573-587.

Wilson, K. S. & von Hippel, P. H. (1995). Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proceedings of the National Academy of Sciences* **92**, 8793-8797.

Wing, J. B., Smart, L., Borrow, R., Findlow, J., Findlow, H., Lees, A., Read, R. C. & Heath, A. W. (2012). Correlation of group C meningococcal conjugate vaccine response with B- and T-lymphocyte activity. *PLoS ONE* **7**, e31160.

Wolfgang, M., Lauer, P., Park, H.-S., Brossay, L., Hébert, J. & Koomey, M. (1998). PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in pilated *Neisseria gonorrhoeae*. *Molecular Microbiology* **29**, 321-330.

Wolfgang, M., Van Putten, J. P. M., Hayes, S. F. & Koomey, M. (1999). The *comP* locus of *Neisseria gonorrhoeae* encodes a type IV prepilin that is dispensable for pilus biogenesis but essential for natural transformation. *Molecular Microbiology* **31**, 1345-1357.

Yager, T. D. & Von Hippel, P. H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. *Biochemistry* **30**, 1097-1118.

Aas, F. E., Løvold, C. & Koomey, M. (2002). An inhibitor of DNA binding and uptake events dictates the proficiency of genetic transformation in *Neisseria gonorrhoeae*: mechanism of action and links to Type IV pilus expression. *Molecular Microbiology* **46**, 1441-1450.

Aas, F. E., Vik, Å., Vedde, J., Koomey, M. & Egge-Jacobsen, W. (2007). *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Molecular Microbiology* **65**, 607-624.

Errata:

List of papers:

I “meningococcal DNA transformation” **changed to** “meningococcal transformation”

IV “polyG” **changed to** “homopolymeric G”

V “Cross-over hotspots in intergenic regions” **changed to** “Intergenic crossover hotspots”

Introduction:

“1.4. Mc genome biology” **changed to** “1.4. Mc genome-related biology”

1.4.3. Line 17 – “in IR DUSs” **changed to** “as IR DUS”

4.4. Line 20 – posses **changed to** impose

“5. References” **changed to** “6. References”

Paper II – **added missing supplementary material** “Table S2”

For the readers’ consideration:

Paper V

The term “Cross-over hotspots/COH” used through out the Introduction and in the Paper V should be read as “breakpoints” or “endpoints” of homologous recombination which are more appropriate terms.

When measuring the allelic distribution of single-nucleotide polymorphisms (SNPs), this is usually referred to as single-nucleotide variants (SNVs).

Structure–function relationships of the competence lipoprotein ComL and SSB in meningococcal transformation

Afsaneh V. Benam,^{1,2†} Emma Lång,^{1,2†} Kristian Alfsnes,^{1,2} Burkhard Fleckenstein,³ Alexander D. Rowe,² Eirik Hovland,¹ Ole Herman Ambur,^{1,2} Stephan A. Frye^{1,2} and Tone Tønjum^{1,2}

Correspondence

Tone Tønjum
tone.tonjum@medisin.uio.no

¹Centre for Molecular Biology and Neuroscience, Institute of Microbiology, University of Oslo, NO-0027 Oslo, Norway

²Centre for Molecular Biology and Neuroscience, Institute of Microbiology, Oslo University Hospital (Rikshospitalet), NO-0027 Oslo, Norway

³Centre for Immune Regulation, Institute of Immunology, University of Oslo, NO-0027 Oslo, Norway

Neisseria meningitidis, the meningococcus, is naturally competent for transformation throughout its growth cycle. The uptake of exogenous DNA into the meningococcus cell during transformation is a multi-step process. Beyond the requirement for type IV pilus expression for efficient transformation, little is known about the neisserial proteins involved in DNA binding, uptake and genome integration. This study aimed to identify and characterize neisserial DNA binding proteins in order to further elucidate the multi-factorial transformation machinery. The meningococcus inner membrane and soluble cell fractions were searched for DNA binding components by employing 1D and 2D gel electrophoresis approaches in combination with a solid-phase overlay assay with DNA substrates. Proteins that bound DNA were identified by MS analysis. In the membrane fraction, multiple components bound DNA, including the neisserial competence lipoprotein ComL. In the soluble fraction, the meningococcus orthologue of the single-stranded DNA binding protein SSB was predominant. The DNA binding activity of the recombinant ComL and SSB proteins purified to homogeneity was verified by electromobility shift assay, and the ComL–DNA interaction was shown to be Mg²⁺-dependent. In 3D models of the meningococcus ComL and SSB predicted structures, potential DNA binding sites were suggested. ComL was found to co-purify with the outer membrane, directly interacting with the secretin PilQ. The combined use of 1D/2D solid-phase overlay assays with MS analysis was a useful strategy for identifying DNA binding components. The ComL DNA binding properties and outer membrane localization suggest that this lipoprotein plays a direct role in neisserial transformation, while neisserial SSB is a DNA binding protein that contributes to the terminal part of the transformation process.

Received 16 October 2010
Revised 18 January 2011
Accepted 16 February 2011

INTRODUCTION

Neisseria meningitidis, or the meningococcus, is a common inhabitant of the mucosal surface of the oro- and nasopharynx in humans. The primary concern regarding meningococcus colonization is the sudden occurrence of systemic meningococcal disease that can occur in

previously healthy individuals (Stephens *et al.*, 2007). The mechanisms that allow some meningococcus strains to disseminate from their local oro-pharyngeal niche and cause acute systemic disease are poorly understood. Most cases of meningococcal disease are caused by clonal complexes of related sequence types (STs), the so-called hyperinvasive lineages (Yazdankhah *et al.*, 2004). These lineages are underrepresented in healthy carriers, and significant numbers of individuals are colonized with carriage isolates belonging to a set of STs that rarely cause disease (Caugant, 2008). Meningococcus cells exhibit abundant antigenic diversity due to frequent recombination, random mutational events, phase variation and high frequencies of horizontal gene transfer (Davidsen &

†These authors have contributed equally to this work.

Abbreviations: DUS, DNA uptake sequence; EMSA, electrophoretic mobility shift assay; ST, sequence type; TPR, tetratricopeptide repeat.

Four supplementary tables, listing strains, plasmids, DNA substrates and PCR primers used in this study, are available with the online version of this paper.

Tønjum, 2006). Natural transformation is the predominant route for exchange of chromosomal DNA between neisserial strains. Unlike other Gram-negative bacteria, meningococcus is naturally competent for transformation throughout its growth cycle (Jysum & Lie, 1965). Transformation in meningococcus is a multi-factorial process that requires the presence of the 12 bp DNA uptake sequence (DUS) in the exogenous DNA (Ambur *et al.*, 2007; Goodman & Scocca, 1988), in addition to type IV pilus expression (Frøholm *et al.*, 1973) and RecA-dependent homologous recombination (Kooimey & Falkow, 1987). Transformation is coupled to the expression of type IV pili in a number of Gram-negative bacteria (Averhoff, 2004; Swanson *et al.*, 1971; Tønjum & Kooimey, 1997). In addition to their role in competence, type IV pili also play a role in adherence (Swanson *et al.*, 1971), twitching motility (Bradley, 1980) and virulence (Caugant, 2008).

Thus, neisserial competence for transformation is dependent on the expression of several pilus-related components, including a number of pilus biogenesis components (Tønjum & Kooimey, 1997) and the minor pilin ComP (Wolfgang *et al.*, 1999). The secretin PilQ is associated with translocation of the pilus fibre across the outer membrane by mediating type IV pilus extrusion and retraction (Tønjum *et al.*, 1998). Neisserial PilQ mutants are non-piliated and non-competent for transformation, and PilQ has previously been shown to bind DNA (Assalkhou *et al.*, 2007). Other components suggested to be involved in neisserial transformation include the competence factors ComE (Chen & Gotschlich, 2001), ComA (Facijs & Meyer, 1993) and ComL (Fussenegger *et al.*, 1996, 1997). However, all of the DNA binding components involved in the machineries which drive meningococcus transformation and other forms of horizontal gene transfer have not yet been identified.

In order to identify and characterize meningococcus proteins that directly bind DNA, we have previously employed cellular fractionation and a solid-phase overlay assay in the form of South-Western analysis in combination with MS analysis (Lång *et al.*, 2009). In general, the procedure employed proved to be a useful strategy for identifying DNA binding components. Here, the resolution of the solid-phase overlay assay was improved by using 2D gel electrophoresis. In the meningococcus membrane and soluble fraction, respectively, the competence lipoprotein ComL and the neisserial orthologue of the single-stranded DNA binding protein SSB were predominant. The observed DNA binding activity of ComL and SSB was verified by electromobility shift analysis. We propose 3D models for the structures of meningococcus ComL and SSB and define their protein-interacting counterparts. Thereby, meningococcus DNA dynamics relevant for horizontal gene transfer and recombination are further elucidated.

METHODS

Bacterial strains, plasmids and growth conditions. Bacterial strains and plasmids employed in this study are listed in

Supplementary Table S1 (available with the online version of this paper). Neisserial strains were grown on blood agar plates in a 5% CO₂ atmosphere at 37 °C for approximately 18 h, while *Escherichia coli* strains were cultivated on Luria-Bertani (LB) plates at 37 °C. Selective antibiotics were added when required.

Cellular fractionation

Enrichment of inner membranes. Isolation of inner membrane proteins from *N. meningitidis* strains and *Neisseria gonorrhoeae* (the gonococcus) strain N400 has previously been described (Lång *et al.*, 2009). Representative fractions were tested for D-lactate dehydrogenase (LDH) activity as a positive control for the enrichment of inner membrane protein (Balasingham *et al.*, 2007).

Enrichment of the soluble fraction. Neisserial cells were resuspended in 1.5 ml PBS, making a bacterial suspension of OD₆₀₀ 0.7. The cells were pelleted at 13 000 g for 10 min, resuspended in 20 µl chloroform and incubated at room temperature for 15 min (Ames *et al.*, 1984; Judd & Porcella, 1993). A 0.1 ml volume of 10 mM Tris/HCl, pH 8.0, was added and the preparation was mixed by vortexing and pelleted by centrifugation at 13 000 g for 20 min at 4 °C. The supernatant containing the soluble fraction was stored at -70 °C.

1D and 2D protein separation and membrane transfer. 1D SDS-PAGE was performed using 10–12% NuPAGE precast gels and sample buffer according to the manufacturer's protocol (Invitrogen), and 2D electrophoresis was performed with Immobiline Drystrips and 10–12% gels according to the manufacturer's recommendations with slight modifications (Amersham GE Healthcare). In brief, isoelectric focusing (IEF) was performed with nonlinear Immobiline DryStrip gel strips of pH 3–10. Prior to analysis, the proteins were solubilized in rehydration buffer [8 M urea, 2 M thiourea, 4% CHAPS, 0.5% Triton X-100, 0.2% DTT, 2% IPG buffer (Amersham GE Healthcare)] overnight before being exposed to ultrasound (Branson 3510) for 10 min. In both 1D and 2D SDS-PAGE, proteins were detected by Coomassie blue staining. The solid-phase overlay experiments, with parallel SDS-PAGE and subsequent MS analysis, were repeated at least three times.

Solid-phase overlay assay for protein–DNA interaction. Protein samples and recombinant full-length ComL were screened for DNA binding activity in a solid-phase overlay assay in the form of a South-Western analysis, which has been described previously (Lång *et al.*, 2009). Protein–DNA interactions were assessed by probing the nitrocellulose membranes (Hybond-C Extra, Amersham GE Healthcare) with biotinylated DNA substrates, with or without DUS (see Supplementary Table S2, available with the online version of this paper). Purified DNA glycosylases Fpg (Tibballs *et al.*, 2009) and MutY, *Taq* polymerase (Sigma) and SSB (Sigma) were used as positive controls, while BSA (Sigma) was used as a negative control.

Protein identification by peptide mass fingerprinting/MALDI-TOF-MS. The DNA binding components detected in the solid-phase overlay assay were identified by MS analysis according to previously described methods (Fleckenstein *et al.*, 2004). For presentation of the results, guidelines provided by Molecular and Cellular Proteomics (http://www.mcponline.org/site/misc/ParisReport_Final.xhtml) were taken into account. In brief, tryptic peptides obtained from in-gel digestion were desalted and concentrated using C18 Empore Extraction Disks (Applied Biosystems) placed in GELoader tips (Eppendorf). The retained peptides were eluted onto a stainless steel target plate (Bruker Daltonics) with a solution containing 70% acetonitrile, 0.1% trifluoroacetic acid and 10 mg α -cyano-4-hydroxycinnamic acid ml⁻¹. After crystallization, the samples were analysed on an Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) operated in the positive reflector mode.

FlexControl v. 2.4 was used for data acquisition, and FlexAnalysis v. 2.2 and BioTools v. 2.2 were used for data evaluation. All MALDI-TOF mass spectra were searched by the web version of Mascot (Matrix Science) against the NCBI, MSDB and Swiss-Prot databases using the following search parameters: one missed cleavage site; fixed modification, carbamidomethylation of cysteine; variable modifications, oxidation of methionine, pyro-glutamate formation of N-terminal glutamine residues; selected taxonomy, *Neisseria meningitidis*; protease, trypsin; peptide mass tolerance ± 0.08 Da.

Bioinformatic analyses and search for signature motifs. The database Uniprot (Boeckmann *et al.*, 2003) provided information on the properties predicted for the DNA binding components identified, whereas DISOPRED2 (Ward *et al.*, 2004) and VSL1 (Obradovic *et al.*, 2005) were employed to assess intrinsic disorder. Searches for functional domains or signature motifs were carried out on the deduced amino acid sequences of meningococcus ComL (AAF41120) and SSB (NP_274471) using the DOLOP (Babu *et al.*, 2006), PROSITE (Hulo *et al.*, 2004) and Pfam databases (Bateman *et al.*, 2004). Based on the ComL and SSB sequences, primary prediction of their secondary structures was performed by using the JPred service (Cuff *et al.*, 1998), revealing a prevalence of alpha-helical elements in ComL, while the presence and location of the transmembrane helices were predicted by MEMSAT3 (Jones *et al.*, 1994). The presence of recognized DNA binding motifs was assessed by using the protscale tool (Gasteiger *et al.*, 2005), while the electrostatic charge was calculated by using the charge program from the EMBOSS package (Rice *et al.*, 2000). In the search for structure-dependent DNA binding motifs, the ComL and SSB sequences were first submitted to the DP-Bind server (Hwang *et al.*, 2007) for prediction of sequence-based DNA binding and subsequently to the PHYRE (Protein Homology/analogy Recognition Engine) server (Bennett-Lovsey *et al.*, 2008) for sequence-based fold recognition and model generation by use of a threading method. Additionally, a homology model for the SSB structure was generated using a SWISS-MODEL (Arnold *et al.*, 2006). Rendered images of 3D structures were generated using PyMOL (DeLano, 2002) and Python Molecular Viewer (Sanner, 1999).

Mutant construction and phenotypic analysis. Meningococcus null mutants corresponding to the DNA binding components identified were constructed using the same strains, plasmids and methods described previously (Supplementary Table S1; Lång *et al.*, 2009) and tested with regard to natural competence for transformation and other pilus-related phenotypes, such as colony morphology and the number of pili. The PCR primers employed in the construction of mutants are listed in Supplementary Table S3 (available with the online version of this paper).

Cloning of the *comL* and *ssb* genes and overexpression of the recombinant proteins. All DNA manipulations were performed according to standard techniques (Maniatis *et al.*, 1982). Full-length (FL) *comL* and *ssb* genes were amplified by PCR from MC58 genomic DNA using the primers listed in Supplementary Table S4 (available with the online version of this paper). The FL *comL* gene was cloned into the expression vector pET28b(+) (Novagen) with a C-terminal 6 \times His-tag, yielding plasmid pAVB1 (Supplementary Table S1). The vector pAVB1 was further used as template in the construction of partial N-terminal and C-terminal ComL constructs using the primers listed in Supplementary Table S4. The FL *ssb* gene was cloned into the vector pQE-30 (Qiagen) with an N-terminal 6 \times His-tag, yielding plasmid pEH1 (Supplementary Table S1). The recombinant proteins were overexpressed in *E. coli* ER2566 (New England Biolabs).

Purification of recombinant ComL and SSB proteins. *E. coli* ER2566 cells overexpressing meningococcus ComL, SSB or ComL

partial protein were grown in LB medium containing 50 μ g kanamycin ml⁻¹ at 37 °C with shaking. The cells were moved to 18 °C at OD₆₀₀ 0.6, induced with 0.5 mM IPTG after 30 min and grown at 18 °C overnight. The cells were harvested by centrifugation at 4000 g for 20 min and frozen at -70 °C. The full-length ComL protein was purified from membrane-enriched fractions and solubilized in 1% n-dodecyl β -maltoside (DDM) (Glycon). Specifically, the cell pellet was resuspended in phosphate buffer (50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0) with the Complete protease inhibitor without EDTA (Roche Applied Science) and benzonase (Merck), and lysed by passing three times through a French press [103 500 kPa, (Thermo Electron)]. Unbroken cells were removed by centrifugation twice at 4000 g for 10 min and the membrane-enriched fraction was collected by ultracentrifugation (150 000 g, 90 min). The membrane pellet was resuspended in a phosphate buffer (pH 8.0) containing 10 mM imidazole and 1% DDM and left to solubilize overnight on a roller at 4 °C. Unsolubilized material was removed by ultracentrifugation (150 000 g, 90 min). The supernatant was added to a Ni-NTA agarose column (Qiagen), washed and eluted with phosphate buffer (pH 8.0) containing 0.1% DDM and increasing amounts of imidazole up to 250 mM. Fractions containing the purified recombinant protein were pooled and dialysed against buffer containing 50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl and 0.1% DDM. The proteins ComL₁₈₋₂₆₇ and SSB were purified from the soluble fraction. In detail, the cell pellets were resuspended in phosphate buffer (50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0) with the Complete protease inhibitor without EDTA (Roche Applied Science) and benzonase (Merck), and lysed by sonication. The lysates were cleared by centrifugation at 16 000 g for 40 min and the supernatant was passed through a Ni-NTA column, and washed and eluted with phosphate buffer (pH 8.0) containing increasing amounts of imidazole up to 250 mM. Fractions containing the purified recombinant proteins were pooled and dialysed against a buffer containing 50 mM NaH₂PO₄, pH 8.0, and 300 mM NaCl.

Solid-phase overlay assay (Far-Western analysis). Protein-protein interactions between ComL, PilQ and PilQ partial proteins were assessed by a solid-phase overlay assay as described previously (Balasingham *et al.*, 2007). Briefly, 1 μ g purified recombinant PilG, PilQ (Frye *et al.*, 2006), PilP, ComL, SSB proteins and BSA were separated by SDS-PAGE and transferred onto nitrocellulose membranes (Hybond-C Extra, Amersham GE Healthcare) in Towbin transfer buffer (25 mM Tris/HCl, 192 mM glycine, 20% methanol, 0.1% SDS, pH 8.3). The membranes were briefly washed twice with renaturing buffer (0.25% gelatin, 0.5% BSA, 0.2% Triton X-100, 10 mM Tris/HCl, 5 mM β -mercaptoethanol, 100 mM NaCl, pH 7.5), and the proteins were renatured by incubation at 4 °C overnight in the same buffer. For the detection of protein-protein binding, the membranes were incubated for 3 h with 1 μ g purified FL or partial proteins of ComL, PilG or PilQ in 10 ml renaturation buffer and washed in Tris-buffered saline (100 mM Tris/HCl, 150 mM NaCl, pH 7.5). Bound ComL, PilG or PilQ was detected with specific rabbit antisera. The PilQ-PilP interaction was used as a positive control (Balasingham *et al.*, 2007) and BSA was used as a negative control. The experiment was repeated at least three times.

Labelling of DNA substrates. Oligonucleotides were end-labelled with [γ -³²P]ATP (Perkin Elmer) using T4 polynucleotide kinase (New England Biolabs) as described by the manufacturer. Labelled oligonucleotides were separated from free nucleotides on 20% non-denaturing PAGE gels and extracted by diffusion into water. Double-stranded labelled substrates were generated by mixing them with an equal molar amount of complementary unlabelled oligomer, heating them to 95 °C for 5 min and slow cooling to room temperature. The concentration of the double-stranded DNA substrate was estimated by dot quantification on agarose plates containing ethidium bromide

as described elsewhere (Maniatis *et al.*, 1982), using unlabelled DNA of known concentration as the standard.

Electrophoretic mobility shift assay (EMSA). For EMSAs, 4.5 fmol labelled DNA was mixed with 4 µl 2.5 × gel shift buffer [25% (w/v) glycerol, 12.5 mM MgCl₂, 0.05 mg BSA ml⁻¹, 2.5 mM DTT] and 5 µl protein diluted in 50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0 and 0.1% DDM in a final volume of 10 µl. The mixture was incubated at 37 °C for 15 min. Electrophoresis was carried out on 6% polyacrylamide gels in Tris/glycine/EDTA buffer (Buratowski & Chodosh, 1996). The gels were dried, exposed to a phosphorImager cassette and scanned in a Typhoon scanner (both from Amersham GE Healthcare). The DNA substrates used in the assay are listed in Supplementary Table S2.

ComL rabbit immunization and antibody production. Rabbit polyclonal antibodies were raised against the ComL_{18–267} protein as described previously (Tønjum *et al.*, 1995). Serum obtained 100 days after immunization was used for ComL detection. Anti-PilQ, anti-PilP, anti-PilQ and anti-PilW sera were produced as described previously (Balasingham *et al.*, 2007; Frye *et al.*, 2006; Tønjum *et al.*, 1995).

Separation of outer and inner membranes by sucrose density gradient. The meningococcus outer and inner membranes were separated by sucrose density gradient as described previously (Balasingham *et al.*, 2007). In brief, meningococcus M1080 cells were processed twice through a French press (103 500 kPa, Thermo Electron). Debris was removed by centrifugation at 10 000 g for 10 min. Sucrose gradient centrifugation was carried out in water with 3 mM EDTA, pH 8.0 (Masson & Holbein, 1983). The sample was transferred onto two layers of sucrose consisting of 3 ml 55% (w/w) sucrose and 4 ml 15% sucrose and centrifuged at 217 000 g and 4 °C for 5 h in an SW40Ti rotor (Beckman). The membrane fraction positioned at the interface was collected and diluted down to 30%

sucrose, applied to a discontinuous sucrose gradient consisting of 3 ml 50, 45, 40 and 35% sucrose and centrifuged in an SW40Ti rotor at 180 000 g and 4 °C for 35 h. After fractionation, 10 µl samples were analysed by SDS-PAGE, followed by Coomassie blue staining and immunoblotting with the ComL-specific antibody.

RESULTS

Search for DNA binding components in the inner membrane and soluble fractions

Proteins in the enriched meningococcus membranes were assessed with regard to their DNA binding activity in a solid-phase overlay assay. Among the multiple DNA binding bands detected in cellular fractions from meningococcus strains representing the major serogroups and one gonococcus strain, a 29 kDa band was the most predominant (Fig. 1a and b). This and four additional reproducibly detected bands were selected for identification. The experiment was performed with ssDNA and dsDNA substrates, with or without DUS. The 29 kDa DNA binding component was shown to interact with both ssDNA and dsDNA substrates, and the DNA binding observed was not enhanced by the presence of DUS.

In order to identify DNA binding proteins in the soluble fraction, chloroform extracts from all neisserial strains tested were subjected to a solid-phase overlay assay. This assessment yielded a predominant 19 kDa DNA binding component in all the isolates examined (Fig. 1c and d). The

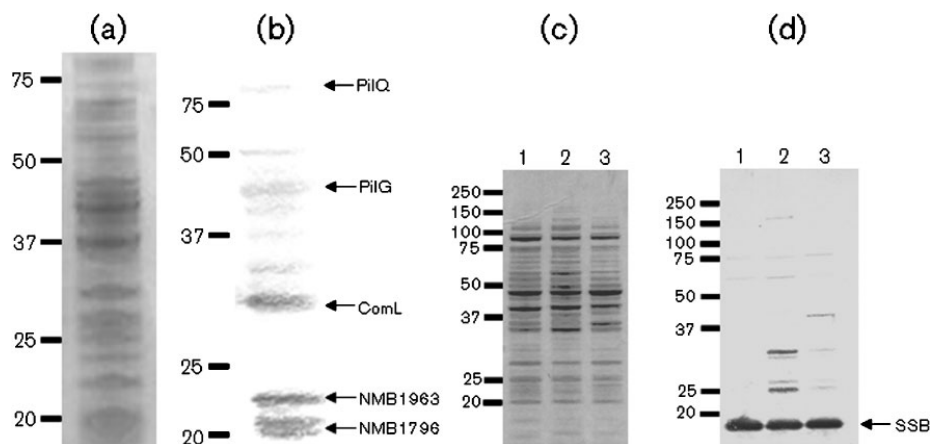


Fig. 1. Identification of DNA binding components by a 1D solid-phase overlay approach. 1D Solid-phase overlay analysis of proteins isolated from the inner membrane and soluble fractions of neisserial strains. (a) 1D Coomassie-blue-stained SDS-PAGE of the inner membrane fraction of meningococcus strain H44/76. (b) Identification of DNA binding components in the inner membrane fraction of meningococcus strain H44/76. (c) 1D Coomassie-blue-stained SDS-PAGE of meningococcus soluble fractions. (d) Identification of SSB as a DNA binding component. Lanes: 1, H44/76; 2, MC58; 3, Z2491. The DNA substrate used in the assays depicted was ssDNA containing the 10 bp DUS. Positions of the size standards (kDa) are shown on the left and arrows on the right indicate the position of the proteins identified by MS analysis.

experiment was performed with ssDNA and dsDNA substrates, with or without DUS. The DNA binding component detected bound to ssDNA and dsDNA, without DUS specificity.

Gel slices containing the six predominant components that reproducibly bound DNA were excised from the parallel Coomassie-blue-stained gels (Fig. 1a and c) and submitted for peptide mass fingerprinting by MALDI-TOF-MS.

Establishment of a 2D solid-phase overlay assay

In order to improve the resolution of protein bands observed in 1D SDS-PAGE, a 2D electrophoresis and subsequent solid-phase overlay assay was performed with the neisserial inner membrane and soluble fractions. Particularly for the inner membrane fraction, the 2D protein separation yielded a higher discrimination between reacting spots of inner membrane fractions in comparison with the 1D separation (Fig. 2). Gel slices containing the six predominant components that reproducibly bound DNA were excised from the parallel Coomassie-blue-stained gels (Fig. 2a and c) and analysed by using MS.

Identification of the DNA binding components by peptide mass fingerprinting/MALDI-TOF-MS

The DNA binding proteins were identified by MALDI-TOF-MS analysis of trypsin-treatment-derived peptides from gel slices (Table 1). The predominant membrane-associated components that repeatedly exhibited DNA binding were identified as the competence lipoprotein ComL (47.9 % sequence coverage, 14 peptides assigned), the pilus biogenesis protein PilG (21.7 % sequence coverage, 11 peptides assigned), NMB1796, a predicted flavoprotein (36.3 % sequence coverage, six peptides assigned), and NMB1963, a periplasmic transport protein (35.7 % sequence coverage, eight peptides assigned). No recognized DNA binding motifs were found in the predicted amino acid sequences of either of these proteins. The additional identification of PilQ in this fraction (22.5 % sequence coverage, 15 peptides assigned) probably either represents contamination by abundant outer membrane proteins due to an incomplete membrane separation or is due to the fact that PilQ directly interacts with inner membrane proteins. PilQ and PilG have been previously found to bind DNA (Assalkhou *et al.*, 2007; Lång *et al.*, 2009).

In the soluble fraction, the predominant 19 kDa DNA binding protein detected was identified as the meningococcus orthologue of the *E. coli* SSB (59.8 % sequence coverage, 10 peptides assigned). No DNA binding motif was identified in the sequence of the neisserial SSB orthologue by bioinformatic analyses, although ssDNA binding activity in the *E. coli* orthologue of SSB has previously been mapped to its C-terminal part, thus solving the crystal structure of the DNA binding domain (Savvides *et al.*, 2004).

Phenotypes of meningococcus null mutants

Meningococcus null mutants corresponding to the DNA binding proteins identified were constructed when possible and examined with regard to their pilus-related phenotypes. The *comL* and *ssb* null mutants, however, were not viable, which corroborates the findings that ComL and SSB homologues are essential in other organisms. The *pilG* and *pilQ* null mutants, which are defective in pilus biogenesis, were non-competent for transformation as described previously (Tønjum *et al.*, 1995, 1998). The NMB1796 and the NMB1963 (Monaco *et al.*, 2006) null mutants were competent for transformation and exhibited wild-type levels of pilus expression (Table 1), suggesting that they are not involved in transformation and that their detection most likely represents a non-specific binding of DNA. Moreover, the identified ORFs encoding proteins that bind DNA were searched for the presence of DUS. DUS located inside ORFs has been shown to be biased towards DNA repair, recombination and replication genes (3R genes), thus indicating a potential role of transforming DNA in genome maintenance (Davidsen *et al.*, 2004). Nevertheless, none of the genes encoding the DNA binding proteins identified contained DUS within their coding sequence.

ComL co-purifies with the meningococcus outer membrane

In order to determine the subcellular localization of ComL, meningococcus outer and inner membranes were separated by sucrose density gradient centrifugation. Fractions from the sucrose gradient were analysed by immunoblotting using antisera against ComL, PilG, PilP, PilQ and PilW (Fig. 3a). PilQ and PilP have previously been shown to reside in the outer and inner membrane, respectively (Balasingham *et al.*, 2007), while PilW resides in the outer membrane, interacting with PilQ (Carbonnelle *et al.*, 2005; Trindade *et al.*, 2008). The main amounts of ComL and PilW peaked with PilQ in the higher-density outer membrane fractions, while PilG and the lipoprotein PilP were concentrated in the inner membrane gradient fractions with very low levels detected in the higher-density fractions, demonstrating that ComL co-purifies with the outer membrane. ComL was shown to be expressed at levels as high as the secretin PilQ (Fig. 3b).

DNA binding properties of recombinant meningococcus ComL and SSB

Native ComL and SSB were identified as neisserial DNA binding proteins by using 1D and 2D solid-phase overlay assays in combination with MS analysis (Figs 1 and 2) (Lång *et al.*, 2009). The DNA binding activity of recombinant ComL and SSB was verified by EMSA (Fig. 4). The ComL DNA binding activity observed was dependent on the presence of Mg^{2+} and was abolished when EMSA buffers containing EDTA or no additive were used (Fig. 5).

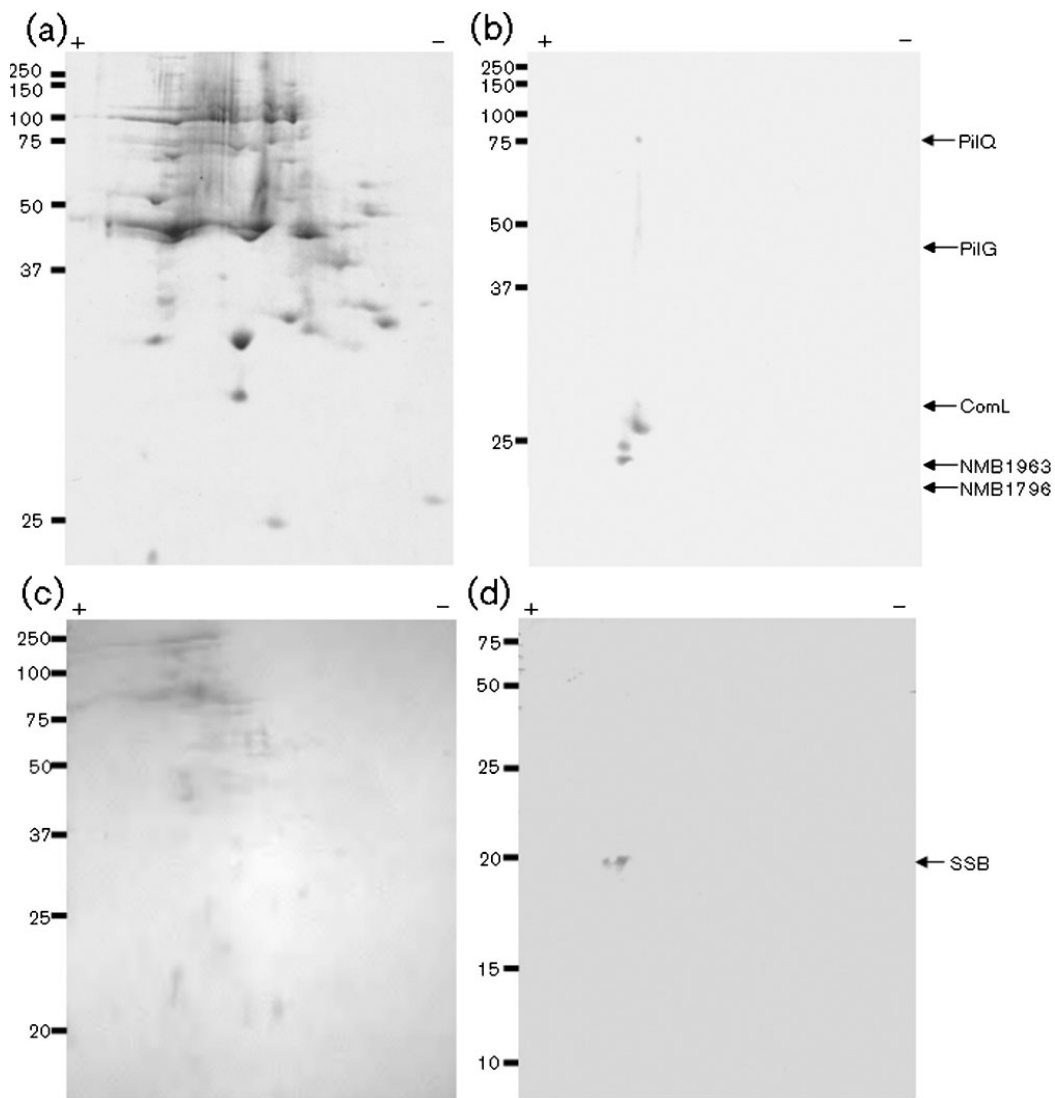


Fig. 2. Identification of DNA binding components by a 2D solid-phase overlay approach and MS analysis. 2D Solid-phase overlay analysis of proteins isolated from the inner membrane and soluble fractions of neisserial strains. (a) 2D Coomassie-blue-stained gel of the inner membrane fraction of meningococcus strain MC58. (b) Identification of DNA binding activity in the inner membrane fraction of MC58 after 2D gel electrophoresis. (c) 2D Coomassie-blue-stained gel of the soluble fraction of MC58. (d) Identification of SSB DNA binding activity in the soluble fraction of MC58 by using the 2D approach. The DNA substrate used in the assays depicted was ssDNA substrate containing the 10 bp DUS. Positions of the size standards (kDa) are shown on the left and arrows on the right indicate the positions of the proteins identified by MS analysis.

Similar results were obtained with all DNA substrates employed (Supplementary Table S2). The DNA binding activities of recombinant ComL and SSB were not DUS-specific (Fig. 4).

ComL and SSB predicted structures

Predicted 3D models for the structure of ComL and SSB were generated using the PHYRE service (Bennett-Lovsey

Table 1. Characteristics of the DNA binding proteins identified by solid-phase overlay assay and MS analysis and phenotypic traits of the corresponding null mutants

nv, Not viable; Wt, wild-type levels.

Protein identified	Predicted size (kDa)	Putative function	Colony morphology*	Extracellular pilus expression	Competence for transformation
Inner membrane fraction					
ComL NMB0703	29	Competence, peptidoglycan-related function	NV	NV	NV
Hypothetical protein NMB1796	20	Predicted flavoprotein	agg +	Wt	Competent
Hypothetical protein NMB1963	21	Periplasmic transport protein	agg +	Wt	Competent
PilG NMB0333	39	Pilus biogenesis	agg −	Absent	Non-competent (Tønjum <i>et al.</i> , 1995)
PilQ NMB1812	80	Pilus biogenesis	agg −	Absent	Non-competent (Frye <i>et al.</i> , 2006)
Soluble fraction					
SSB† NMB1460	19	ssDNA-binding protein	NV	NV	NV

*The colony morphology is described as agglutinating (agg +) and non-agglutinating (agg −).

†SSB was the single predominant component identified in the soluble fraction.

et al., 2008). The sequence identity of ComL relative to known structures was below the 30% level required to generate a confident homology model, with the best results returning between 11 and 18% sequence identity. However, the *e*-values of these alignments were very low ($<1e^{-19}$), which allowed for the generation of a hypothetical structure by threading. The most compatible structures returned by the server were characterized due to the presence of a tetratricopeptide repeat (TPR) structure; these findings included *Pseudomonas aeruginosa* PilF (Kim *et al.*, 2006), which was also the component closest in length to ComL. Since sequence-based DNA binding predictions with regard to ComL were uniformly negative, the threading model structure was used as a basis to investigate the possible distribution of positive charges on the surface, thereby yielding a possible explanation for the DNA binding observed. A 3D model structure was generated and the charge on the molecular surface was calculated by using the Adaptive Poisson–Boltzmann Solver (Fig. 6) (Baker *et al.*, 2001). Charge distribution on the ComL surface is shown for a representative range of orientations, demonstrating that there are several regions of positive charge (blue), which may function in specific or non-specific DNA binding.

In contrast, neisserial SSB had 33.6% sequence identity with a known 3D crystal structure from *Thermus aquaticus* (Fedorov *et al.*, 2006), which allowed the automated generation of a homology model by means of the SWISS-MODEL server pipeline (Arnold *et al.*, 2006). An additional structural SSB homologue from *E. coli* has been crystallized in its multimeric form, bound to ssDNA (Raghunathan *et al.*, 2000). The 3D model structure for neisserial SSB has been substituted as a dimer into this structure in order to illustrate the predicted meningococcus SSB mode of binding for ssDNA (Fig. 7).

ComL directly interacts with N-terminal PilQ

A solid-phase overlay assay in the form of a Far-Western analysis was employed to determine if there was interaction between the ComL and PilQ proteins (Fig. 8). This revealed that the ComL and N-terminal PilQ proteins directly interact *in vitro*.

DISCUSSION

In this study, we searched cellular fractions from a representative panel of neisserial strains for DNA binding components, using 1D and 2D electrophoresis combined with a solid-phase overlay approach. The DNA binding proteins were identified by MS analysis. Enrichment of the membrane and the soluble fractions were important steps contributing to the generation of target solutions with reduced complexity, as compared with meningococcus whole-cell extracts, for the identification of DNA binding proteins. The 2D solid-phase overlay strategy further increased the spatial resolution of the proteins in comparison with the 1D approach. Employing peptide mass fingerprinting by MALDI-TOF-MS enabled protein identification from small amounts present in the 1D/2D gel spots which exhibited DNA binding.

The predominant membrane components exhibiting DNA binding activity were identified as ComL and PilG, as previously reported (Lång *et al.*, 2009). In this context, we wanted to characterize ComL further. In EMSA, ComL was shown to bind DNA in a Mg^{2+} -dependent manner, indicating that Mg^{2+} facilitates the direct binding of DNA (Fig. 5). A hypothetical 3D structure for ComL was generated and the charge distribution on the ComL molecular surface was predicted, suggesting that there are

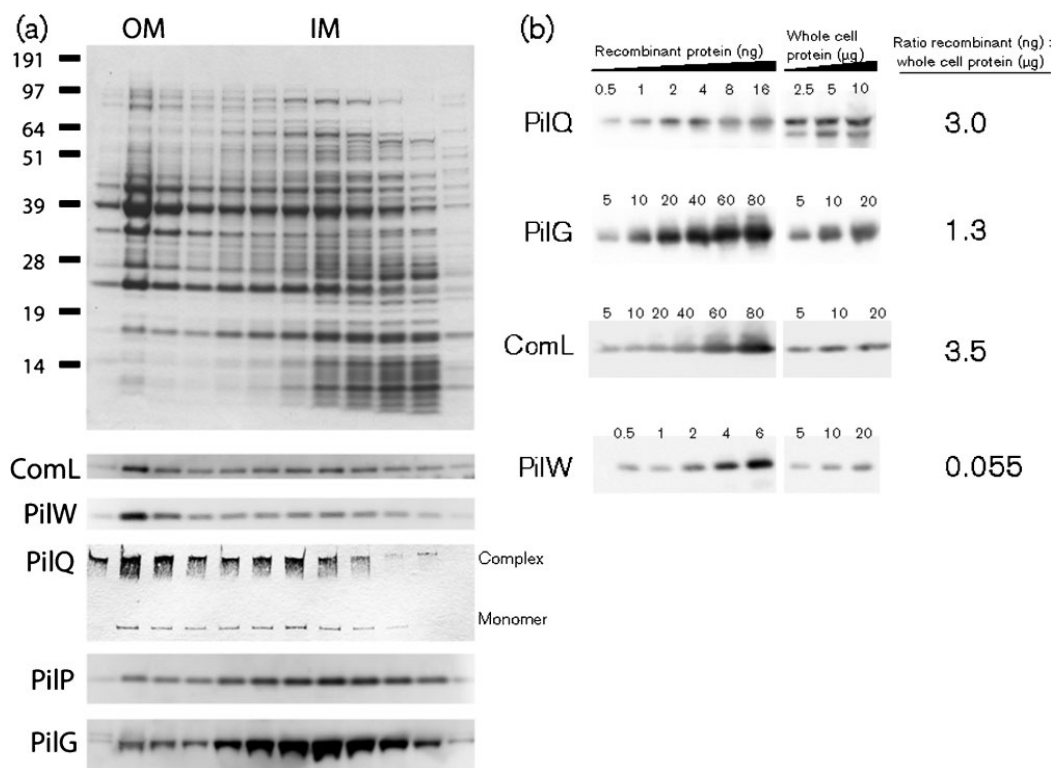


Fig. 3. ComL co-purifies with the outer membrane and is expressed at high levels. (a) Analysis of cellular fractions after separation of outer and inner membranes from *N. meningitidis* M1080 by sucrose-gradient centrifugation. Samples from each fraction were separated by SDS-PAGE and stained with Coomassie blue (top panel) or detected by immunoblotting using antibodies against ComL, PilG, PilP, PilQ and PilW (lower panels). (b) Quantitative immunoblotting of a defined amount of *N. meningitidis* M1080 cellular suspension using the same antibodies as in (a).

several regions of positive charge, which may function as specific or non-specific DNA binding regions. ComL was also shown to have structural homology with the *P. aeruginosa* pilus biogenesis protein PilF (Kim *et al.*, 2006), which is homologous to the neisserial outer membrane lipoprotein PilW engaged in pilus biogenesis (Carbannelle *et al.*, 2005; Trindade *et al.*, 2008).

The neisserial ComL protein has been suggested to contribute to DNA uptake by cleavage of the peptidoglycan layer during transformation (Fussenegger *et al.*, 1996, 1997). The gonococcal *comL* gene exists in a single copy, which is transcribed in the opposite direction to the neighbouring *comA* gene and encodes a periplasmic lipoprotein with a relatively high theoretical pI (9.03). The *comL* and *comA* gene pair has a common DUS-containing transcriptional terminator in an appropriate position for joint use (Jose *et al.*, 2003). The ComL orthologue in *E. coli*, YfiO, is anchored to the outer

membrane and is a member of the β -barrel assembly machinery (the BAM complex) (Wu *et al.*, 2005). Recently, ComL was referred to as an orthologue of BamD, which suggests that neisserial ComL is associated with the outer membrane and is involved in outer membrane protein biogenesis (Knowles *et al.*, 2009). This notion was supported by our membrane separation (Fig. 3a). Orthologues of ComL are conserved among both neisserial and other Gram-negative bacterial species (Fussenegger *et al.*, 1996; Malinverni *et al.*, 2006). Bioinformatic inferences suggest that the *comL* gene encodes TPRs, and that the resulting structure presents several large regions of positive charge, which may act as either ssDNA or dsDNA binding sites. Still, it is not possible to determine from these structural predictions whether the apparent DNA binding capacity is related to the TPRs. TPRs have been predicted for neisserial PilW and *Pseudomonas* PilF, and have been suggested to serve a functional role in PilQ stabilization (Koo *et al.*, 2008; Trindade *et al.*, 2008). Thus,

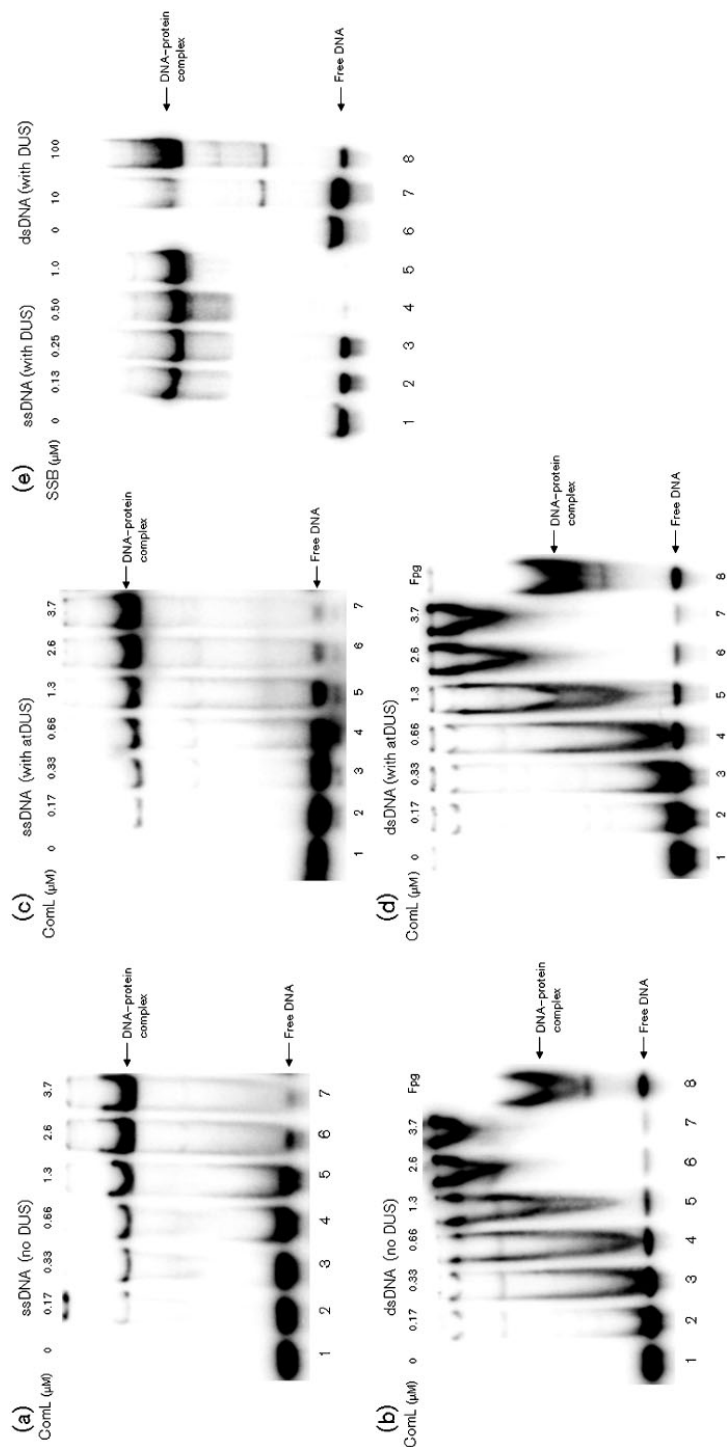


Fig. 4. *N. meningitidis* recombinant ComL and SSB exhibit DNA binding activity. The DNA binding abilities of purified recombinant full-length ComL and SSB proteins were assessed by EMSA. (a) ComL binds to ssDNA substrate without the DUS. A protein concentration between 0.66 and 1.3 μM is required to shift 50% of the ssDNA substrate, indicating an apparent dissociation constant of 0.66–1.3 μM . (b) ComL binds to dsDNA substrate without DUS. A protein concentration between 0.33 and 0.66 μM is required to shift 50% of the dsDNA substrate, indicating an apparent dissociation constant of 0.33–0.66 μM . A 50 ng sample of recombinant meningococcus Fpg was used as a positive control. (c, d) The same results were obtained with ssDNA and dsDNA substrates with the 12 bp DUS. (e) SSB binds to both ssDNA and dsDNA, though with a strong preference for ssDNA. A protein concentration between 0.25 and 0.5 μM is required to shift 50% of the ssDNA substrate, and between 10 and 100 μM is required to shift 50% of the dsDNA substrate. This indicates an approximately 150-fold higher affinity for ssDNA than for dsDNA. atDUS, 12 bp DUS.

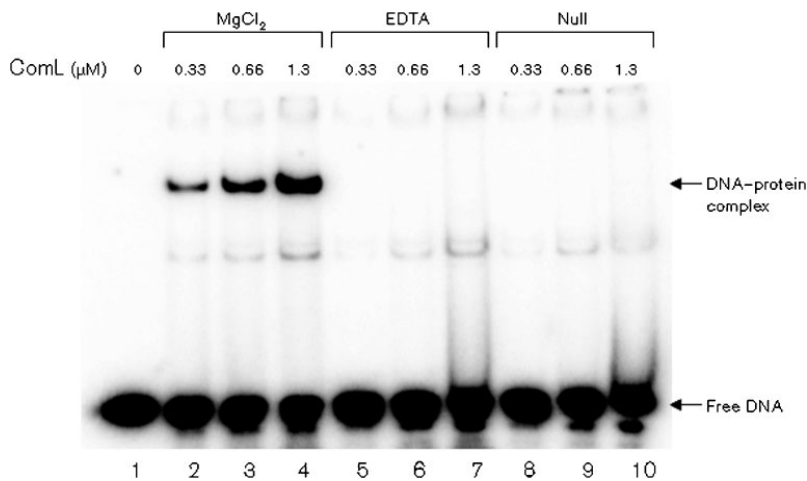


Fig. 5. *N. meningitidis* recombinant ComL DNA binding activity is Mg²⁺-dependent. A panel of EMSA buffers was used to assess the DNA binding activity of the purified full-length ComL protein. A protein–DNA complex shift was observed only with buffer containing MgCl₂. This shift was abolished when buffers containing EDTA or no additives were used.

as previously suggested, the TPRs predicted could be of functional importance in mediating protein–protein interactions between ComL and other BAM proteins (D’Andrea & Regan, 2003; Knowles *et al.*, 2009). No potential DNA–protein interactions have been reported for TPR-containing proteins (D’Andrea & Regan, 2003), though the predicted positively charged regions of ComL may interact with DNA in a non-specific manner. Site-directed point

mutations will enable the elucidation of the potential role of TPRs in ComL-mediated DNA binding or protein–protein interactions.

A significant decrease in the transformation rate of a gonococcal *comL* mutant has been documented, suggesting a role of the lipoprotein ComL in the neisserial transformation process in interaction with the peptidoglycan layer

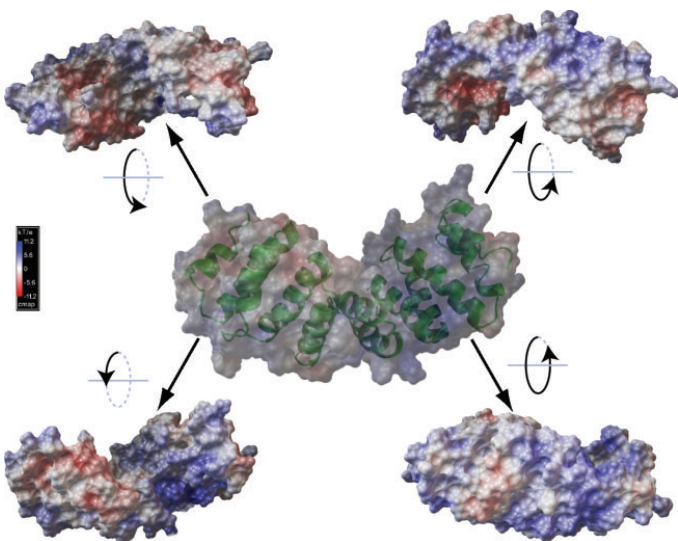


Fig. 6. *N. meningitidis* ComL 3D model. The *N. meningitidis* MC58 ComL hypothetical 3D structure generated by the PHYRE threading server is shown in cartoon form (centre) with the concurrent surface representation overlaid. The charge distribution, calculated using APBS, is shown on the surface (blue=positive, red=negative). The ComL surface is shown in four further images (in corners) rotated progressively around a horizontal axis in order to give a complete overview of the charge distribution. The large areas of positive charge may indicate that they play a role in DNA binding.

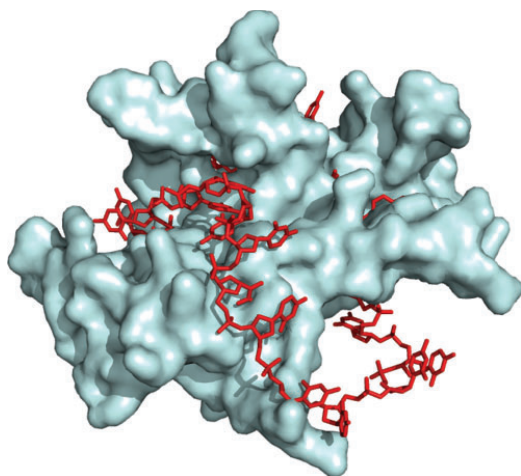


Fig. 7. *N. meningitidis* SSB 3D model. The *N. meningitidis* MC58 SSB predicted 3D structure modelled by homology using the SWISS-MODEL service is shown here as a dimer, substituted into the crystal structure for *E. coli* SSB bound to ssDNA. This illustrates a functional mode for ssDNA binding by SSB.

(Fussenegger *et al.*, 1996). Moreover, neisserial ComL has been suggested to be involved in the folding of outer membrane proteins (Knowles *et al.*, 2009). The true function of ComL is difficult to assess since most mutants of this essential component are lethal. It is of note that Fussenegger and co-workers managed to generate a viable gonococcal *comL* mutant, expressing a truncated version of the protein, therefore indicating the importance of an intact N terminus in ComL protein expression and function (Fussenegger *et al.*, 1996). Meningococcal *comL* null mutants were also non-viable, indicating that ComL is essential in neisserial species (Table 1). The *pilQ* and *pilG* null mutants, defective in pilus biogenesis, were non-competent for transformation. For these two components, the biological significance of their DNA binding capabilities is complicated by the fact that they participate in type IV pilus biogenesis, which in turn is required for competence. Thus, it is a conundrum as to whether the lack of competence in these mutants is due to a defect in their direct binding of DNA or whether this lack is indirect through pilus biogenesis.

SSB, the predominant DNA binding protein detected in the soluble neisserial cell fractions, is involved in processing ssDNA intermediates during DNA replication, repair and recombination in *E. coli* (Shereda *et al.*, 2008). Furthermore, SSB proteins are conserved, serve critical functions in genome maintenance and are indispensable for cell survival among both prokaryotes and eukaryotes (Fanning *et al.*, 2006; Shereda *et al.*, 2008; Zou *et al.*, 2006). In our hands, neisserial SSB was found to bind both ssDNA and

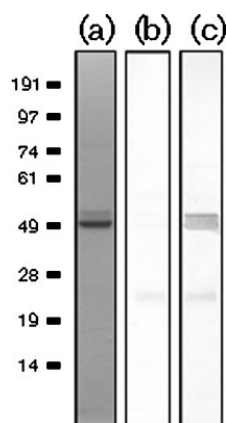


Fig. 8. The lipoprotein ComL directly interacts with PilQ. Solid-phase overlay interaction study between ComL_{18–267} and PilQ_{25–354} on the blot and overlay with ComL_{18–267}, and detection with ComL antibody. (a) Coomassie blue staining of PilQ_{25–354}. (b) Negative control with PilQ_{25–354} protein on the membrane probed with ComL antibody alone. (c) PilQ_{25–354} protein on the membrane, probed with ComL antibody after an overlay of ComL_{18–267}. A stronger staining for PilQ_{25–354} and its degradation product is seen when an overlay of ComL_{18–267} is added before detection with antibody, indicating an interaction between PilQ_{25–354} and ComL_{18–267}.

dsDNA, which is consistent with findings on SSB in other prokaryotes, in which SSB either binds to ssDNA or intercalates itself into dsDNA, thereby disrupting it (Makhov & Griffith, 2006; Makhov *et al.*, 2009; Mapelli *et al.*, 2005). Previously, the genome of *Bacillus subtilis* was shown to comprise two paralogous SSB genes, *ssb* and *ywpH*. Interestingly, the proteins encoded by *ssb* and *ywpH* have distinctive expression patterns, with SSB being essential for cell survival, while YwpH is required for natural transformation (Lindner *et al.*, 2004). In the *ywpH* null mutant, the transformation rate was reduced fivefold, whereas the *ssb* null mutant was not viable (Lindner *et al.*, 2004; Ogura *et al.*, 2002). In *Haemophilus influenzae* RD KW20, gene expression analyses have revealed that the SSB orthologue HI0250 is induced 3.4-fold during the competent state (Redfield *et al.*, 2005). Based on the broad conservation of SSB functions and the documented role of SSB in transformation, we suggest that neisserial SSB might also have a functional role in transformation. As in *E. coli* (Shereda *et al.*, 2008), meningococcus *ssb* null mutants were non-viable, and meningococcus SSB phenotypes directly associated with transformation could therefore not be assessed (Table 1).

The purpose of this study was to identify and characterize neisserial DNA binding proteins and assess their potential relevance for transformation, putting a special emphasis on the characterization of ComL and SSB and their DNA binding properties. The neisserial DUS has been shown to

mediate selective uptake of DNA through transformation; thus the search for DUS-specific DNA binding components was a priority. However, this endeavour turned out to be a difficult task since none of the DNA binding proteins identified bound DNA in a DUS-specific manner, nor contained the DUS sequence within their ORFs. The difficulties and challenges in identifying a potential DUS-specific binding protein are multiple, since functional and technical obstacles obscure the hunt for an unknown or putative component. If it does exist, the cellular location of DUS selectivity is not yet known. This search for DNA binding candidates targeted the inner membrane and cytoplasm, and a DUS-specific protein in these fractions and the neisserial outer membrane has not yet been identified. This negative search result could well be due to technical limitations, in that the *in vitro* conditions in the solid-phase overlay approach employed do not reflect the DNA binding that goes on *in vivo*. The solid-phase overlay assay also has limitations in the form of the level of protein expression, protein folding and renaturation abilities after SDS-PAGE. Thereby, not all DNA binding candidates will be detected using this method, and DUS-specific DNA binding might be of such a subtle and transient nature that it is not detected in this assay. Yet, the reproducible identification of PilG, ComL and PilQ DNA binding by this approach (Lång *et al.*, 2009), in addition to other independent assays (Assalkhou *et al.*, 2007), is strong evidence that the method is indeed valid for detecting a number of DNA binding proteins in general.

This study strengthens previous findings on potential direct roles for ComL, PilG and PilQ in transformation. Additional studies are warranted to provide new insights into the functional relationships between these and other proteins involved in the transformation process. Characterization of the physical interactions between ComL and SSB with DNA, in addition to other proteins, will contribute to a better understanding of how transforming DNA is processed in meningococcus cells and will further elucidate the neisserial DNA uptake and genome integration process.

ACKNOWLEDGEMENTS

The Medical Research Curriculum and the Institute of Medical Basic Sciences at the University of Oslo are greatly acknowledged for their support to A. V. B. We are grateful to Håvard Homberset for excellent technical assistance and invaluable discussions and to Marit Jørgensen for protein identification at the Proteomics Core Facility at Oslo University Hospital (Rikshospitalet)/University of Oslo. This work was supported by a Centre of Excellence grant from the Research Council of Norway to the Centre for Molecular Biology and Neuroscience (CMBN) and the FUGE platform Consortium for Microbial Science Technology (CAMST).

REFERENCES

Ambur, O. H., Frye, S. A. & Tønnum, T. (2007). New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators. *J Bacteriol* **189**, 2077–2085.

Ames, G. F., Prody, C. & Kustu, S. (1984). Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J Bacteriol* **160**, 1181–1183.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201.

Assalkhou, R., Balasingham, S., Collins, R. F., Frye, S. A., Davidsen, T., Benam, A. V., Bjørås, M., Derrick, J. P. & Tønnum, T. (2007). The outer membrane secretin PilQ from *Neisseria meningitidis* binds DNA. *Microbiology* **153**, 1593–1603.

Averhoff, B. (2004). DNA transport and natural transformation in mesophilic and thermophilic bacteria. *J Bioenerg Biomembr* **36**, 25–33.

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L. & Sankaran, K. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* **188**, 2761–2773.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037–10041.

Balasingham, S. V., Collins, R. F., Assalkhou, R., Homberset, H., Frye, S. A., Derrick, J. P. & Tønnum, T. (2007). Interactions between the lipoprotein PilP and the secretin PilQ in *Neisseria meningitidis*. *J Bacteriol* **189**, 5716–5727.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S. & other authors (2004). The Pfam protein families database. *Nucleic Acids Res* **32** (Database issue), D138–D141.

Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. & Kelley, L. A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611–625.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C. & other authors (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* **31**, 365–370.

Bradley, D. E. (1980). A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. *Can J Microbiol* **26**, 146–154.

Buratowski, S. & Chodosh, L. A. (1996). Mobility shift DNA-binding assay using gel electrophoresis. In *Current Protocols in Molecular Biology*, Chapter 12, 12.2.1–12.2.11. Edited by F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith & K. Struhl. New York: Wiley.

Carbonnelle, E., Hélaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. (2005). Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol Microbiol* **55**, 54–64.

Caugant, D. A. (2008). Genetics and evolution of *Neisseria meningitidis*: importance for the epidemiology of meningococcal disease. *Infect Genet Evol* **8**, 558–565.

Chen, I. & Gotschlich, E. C. (2001). ComE, a competence protein from *Neisseria gonorrhoeae* with DNA-binding activity. *J Bacteriol* **183**, 3160–3168.

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. & Barton, G. J. (1998). JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892–893.

D'Andrea, L. D. & Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem Sci* **28**, 655–662.

Davidsen, T. & Tønnum, T. (2006). Meningococcal genome dynamics. *Nat Rev Microbiol* **4**, 11–22.

Davidsen, T., Rødland, E. A., Lagesen, K., Seeberg, E., Rognes, T. & Tønnum, T. (2004). Biased distribution of DNA uptake sequences towards genome maintenance genes. *Nucleic Acids Res* **32**, 1050–1058.

- DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. San Carlos, CA: DeLano Scientific LLC. <http://www.pymol.org>
- Facius, D. & Meyer, T. F. (1993). A novel determinant (*comA*) essential for natural transformation competence in *Neisseria gonorrhoeae* and the effect of a *comA* defect on pilin variation. *Mol Microbiol* 10, 699–712.
- Fanning, E., Klimovich, V. & Nager, A. R. (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* 34, 4126–4137.
- Fedorov, R., Witte, G., Urbanke, C., Manstein, D. J. & Curth, U. (2006). 3D structure of *Thermus aquaticus* single-stranded DNA-binding protein gives insight into the functioning of SSB proteins. *Nucleic Acids Res* 34, 6708–6717.
- Fleckenstein, B., Qiao, S. W., Larsen, M. R., Jung, G., Roepstorff, P. & Sollid, L. M. (2004). Molecular characterization of covalent complexes between tissue transglutaminase and gliadin peptides. *J Biol Chem* 279, 17607–17616.
- Frøholm, L. O., Jysum, K. & Bøvre, K. (1973). Electron microscopical and cultural features of *Neisseria meningitidis* competence variants. *Acta Pathol Microbiol Scand B Microbiol Immunol* 81, 525–537.
- Frye, S. A., Assalkhou, R., Collins, R. F., Ford, R. C., Petersson, C., Derrick, J. P. & Tønjum, T. (2006). Topology of the outer-membrane secretin PilQ from *Neisseria meningitidis*. *Microbiology* 152, 3751–3764.
- Fussenegger, M., Facius, D., Meier, J. & Meyer, T. F. (1996). A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of *Neisseria gonorrhoeae*. *Mol Microbiol* 19, 1095–1105.
- Fussenegger, M., Rudel, T., Barten, R., Ryll, R. & Meyer, T. F. (1997). Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae* – a review. *Gene* 192, 125–134.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, pp. 571–607. Edited by J. M. Walker. Totowa, NJ: Humana Press.
- Goodman, S. D. & Scocca, J. J. (1988). Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proc Natl Acad Sci U S A* 85, 6982–6986.
- Hulo, N., Sigrist, C. J., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P. & Bairoch, A. (2004). Recent improvements to the PROSITE database. *Nucleic Acids Res* 32 (Database issue), D134–D137.
- Hwang, S., Gou, Z. & Kuznetsov, I. B. (2007). DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins. *Bioinformatics* 23, 634–636.
- Jones, D. T., Taylor, W. R. & Thornton, J. M. (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* 33, 3038–3049.
- Jose, J., Otto, G. W. & Meyer, T. F. (2003). The integration site of the *iga* gene in commensal *Neisseria* sp. *Mol Genet Genomics* 269, 197–204.
- Judd, R. C. & Porcella, S. F. (1993). Isolation of the periplasm of *Neisseria gonorrhoeae*. *Mol Microbiol* 10, 567–574.
- Jysum, K. & Lie, S. (1965). Genetic factors determining competence in transformation of *Neisseria meningitidis*. 1. A permanent loss of competence. *Acta Pathol Microbiol Scand* 63, 306–316.
- Kim, K., Oh, J., Han, D., Kim, E. E., Lee, B. & Kim, Y. (2006). Crystal structure of PilF: functional implication in the type 4 pilus biogenesis in *Pseudomonas aeruginosa*. *Biochem Biophys Res Commun* 340, 1028–1038.
- Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* 7, 206–214.
- Koo, J., Tammam, S., Ku, S. Y., Sampaleanu, L. M., Burrows, L. L. & Howell, P. L. (2008). PilF is an outer membrane lipoprotein required for multimerization and localization of the *Pseudomonas aeruginosa* Type IV pilus secretin. *J Bacteriol* 190, 6961–6969.
- Koomey, J. M. & Falkow, S. (1987). Cloning of the *recA* gene of *Neisseria gonorrhoeae* and construction of gonococcal *recA* mutants. *J Bacteriol* 169, 790–795.
- Lång, E., Haugen, K., Fleckenstein, B., Homberset, H., Frye, S. A., Ambur, O. H. & Tønjum, T. (2009). Identification of neisserial DNA binding components. *Microbiology* 155, 852–862.
- Lindner, C., Nijland, R., van Hartskamp, M., Bron, S., Hamoen, L. W. & Kuipers, O. P. (2004). Differential expression of two paralogous genes of *Bacillus subtilis* encoding single-stranded DNA binding protein. *J Bacteriol* 186, 1097–1105.
- Makhov, A. M. & Griffith, J. D. (2006). Visualization of the annealing of complementary single-stranded DNA catalyzed by the herpes simplex virus type 1 ICP8 SSB/recombinase. *J Mol Biol* 355, 911–922.
- Makhov, A. M., Sen, A., Yu, X., Simon, M. N., Griffith, J. D. & Egelman, E. H. (2009). The bipolar filaments formed by herpes simplex virus type 1 SSB/recombination protein (ICP8) suggest a mechanism for DNA annealing. *J Mol Biol* 386, 273–279.
- Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* 61, 151–164.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Mapelli, M., Panjikar, S. & Tucker, P. A. (2005). The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J Biol Chem* 280, 2990–2997.
- Masson, L. & Holbein, B. E. (1983). Physiology of sialic acid capsular polysaccharide synthesis in serogroup B *Neisseria meningitidis*. *J Bacteriol* 154, 728–736.
- Monaco, C., Talà, A., Spinosa, M. R., Progida, C., De Nitto, E., Gaballo, A., Bruni, C. B., Buccì, C. & Alifano, P. (2006). Identification of a meningococcal L-glutamate ABC transporter operon essential for growth in low-sodium environments. *Infect Immun* 74, 1725–1740.
- Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P. & Dunker, A. K. (2005). Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* 61 (Suppl. 7), 176–182.
- Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y. & Tanaka, T. (2002). Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol* 184, 2344–2351.
- Raghunathan, S., Kozlov, A. G., Lohman, T. M. & Waksman, G. (2000). Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat Struct Biol* 7, 648–652.
- Redfield, R. J., Cameron, A. D., Qian, Q., Hinds, J., Ali, T. R., Kroll, J. S. & Langford, P. R. (2005). A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. *J Mol Biol* 347, 735–747.
- Rice, P., Longden, I. & Bleasby, A. (2000). EMBOS: the European Molecular Biology Open Software Suite. *Trends Genet* 16, 276–277.
- Sanner, M. F. (1999). Python: a programming language for software integration and development. *J Mol Graph Model* 17, 57–61.
- Savvides, S. N., Raghunathan, S., Fütterer, K., Kozlov, A. G., Lohman, T. M. & Waksman, G. (2004). The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci* 13, 1942–1947.

- Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. & Keck, J. L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **43**, 289–318.
- Stephens, D. S., Greenwood, B. & Brandtzaeg, P. (2007). Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*. *Lancet* **369**, 2196–2210.
- Swanson, J., Kraus, S. J. & Gotschlich, E. C. (1971). Studies on gonococcus infection. I. Pili and zones of adhesion: their relation to gonococcal growth patterns. *J Exp Med* **134**, 886–906.
- Tibballs, K. L., Ambur, O. H., Alfsnes, K., Homberst, H., Frye, S. A., Davidsen, T. & Tønnum, T. (2009). Characterization of the meningococcal DNA glycosylase Fpg involved in base excision repair. *BMC Microbiol* **9**, 7.
- Tønnum, T. & Koomey, M. (1997). The pilus colonization factor of pathogenic neisserial species: organelle biogenesis and structure/function relationships – a review. *Gene* **192**, 155–163.
- Tønnum, T., Freitag, N. E., Namork, E. & Koomey, M. (1995). Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. *Mol Microbiol* **16**, 451–464.
- Tønnum, T., Caugant, D. A., Dunham, S. A. & Koomey, M. (1998). Structure and function of repetitive sequence elements associated with a highly polymorphic domain of the *Neisseria meningitidis* PilQ protein. *Mol Microbiol* **29**, 111–124.
- Trindade, M. B., Job, V., Contreras-Martel, C., Pelicic, V. & Dessen, A. (2008). Structure of a widely conserved type IV pilus biogenesis factor that affects the stability of secretin multimers. *J Mol Biol* **378**, 1031–1039.
- Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. & Jones, D. T. (2004). The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138–2139.
- Wolfgang, M., van Putten, J. P., Hayes, S. F. & Koomey, M. (1999). The *comP* locus of *Neisseria gonorrhoeae* encodes a type IV prepilin that is dispensable for pilus biogenesis but essential for natural transformation. *Mol Microbiol* **31**, 1345–1357.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235–245.
- Yazdankhah, S. P., Kriz, P., Tzanakaki, G., Kremastinou, J., Kalmusova, J., Musilek, M., Alvestad, T., Jolley, K. A., Wilson, D. J. & other authors (2004). Distribution of serogroups and genotypes among disease-associated and carried isolates of *Neisseria meningitidis* from the Czech Republic, Greece, and Norway. *J Clin Microbiol* **42**, 5146–5153.
- Zou, Y., Liu, Y., Wu, X. & Shell, S. M. (2006). Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. *J Cell Physiol* **208**, 267–273.

Edited by: P. van der Ley

Supplementary Table S1. Bacterial strains and plasmids employed in this study

Strain or plasmid		Relevant characteristic	Reference or source
<i>N. meningitidis</i> strains			
Z2491	Serogroup A, isolated in	Gambia 1983	Crowe <i>et al.</i> (1989)
MC58	Serogroup B, isolated in the	UK 1983	McGuinness <i>et al.</i> (1991)
H44/76	Serogroup B, isolated in	Norway 1976	Holten (1979)
M1080	Serogroup B, isolated in the	USA, 1969	Frasch & Chapman (1972)
8013	Serogroup C		Caugant <i>et al.</i> (1986)
H44/76 ΔPilG	<i>pilG::mTnErm</i> transposon insertion		Tønjum <i>et al.</i> (1995)
H44/76 ΔPilQ	H44/76 strain with mTn::cm#21 in <i>pilQ</i>		Tønjum <i>et al.</i> (1998)
<i>N. gonorrhoeae</i> strain			
N400	MS11 subclone, <i>recA6(tetM)*</i>		Tønjum <i>et al.</i> (1995)
<i>E. coli</i> strains			
ER2566	Expression strain with a chromosomal copy of the T7 RNA polymerase gene		New England Biolabs
XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI^q</i> ΔM15 Tn10Tet']		Stratagene
Vectors			
pBSK+	General cloning vector, Amp ^r		Stratagene
pET28b(+)	Expression vector based on a T7-promoter-driven system, His tag, Kan ^r		Novagen
pQE-30	Expression vector based on a T5-promoter-driven system, His tag, Amp ^r		Qiagen
pAVB1	pET28b harbouring <i>comL</i> from MC58		This study
pHH10	pET28b harbouring ComL _{18–267} insert		This study
pHH11	pET28b harbouring ComL _{18–140} insert		This study
pHH12	pET28b harbouring ComL _{139–267} insert		This study
pEH1	pEQ-30 harbouring <i>ssb</i> from MC58		This study

**recA6* is an IPTG-inducible allele of *recA*.

Supplementary Table S2. DNA substrates employed in this study

Substrate	Sequence (5'→3')*	DUS length
ssDNA		
T ₁	CAACAACAACAACAGCC GTCTGA ACCAAATTCAGACGGCAACAACAACAACA	+, 10 bp
T ₃	CAACAACAACAACAGGCCTGTCATCCAAACTGACAGGCCAACAACAACAACA	–
HH7†	AACAACAACAAAT GCCGTCTGA ACCAACATGCCGTCTGAAAACAACAACAAC	+, 12 bp
HH9	GTTGTTGTTGTTAT GCCGTCTGA AGTTGGATGCCGTCTGAATTGTTGTTGTT	+, 12 bp
HH10†	AACAACAACAAAAGGCCTGTCATCCAACTTCCGGACAGTAAACAACAACAAC	–
dsDNA		
T ₁ T ₂	CAACAACAACAACAGCC GTCTGA ACCAAATTCAGACGGCAACAACAACAACA TGTGTTGTTGTTGCCGTCTGAATTTGG TTCAGACGGC CTGTTGTTGTTGTTG	+, 10 bp
HH7, HH8†	AACAACAACAAAT GCCGTCTGA ACCAACATGCCGTCTGAAAACAACAACAAC GTTGTTGTTGTTTTCAGACGGCATGTTGG TTCAGACGGC ATTGTTGTTGTT	+, 12 bp
HH10, HH11†	AACAACAACAAAAGGCCTGTCATCCAACTTCCGGACAGTAAACAACAACAAC GTTGTTGTTGTTTACTGTCCGGAAGTTGGATGACAGGCCTTTTGTTGTTGTT	–

*The DNA uptake sequence (DUS) is indicated by bold type.

†DNA substrates used in electromobility shift analysis.

Supplementary Table S3. Primers employed in construction of mutants

Primer	Sequence (5'→3')*	Restriction sites
<i>comL</i> (NMB0703)		
KH45	GCGCTAGCTCTTGGGTAATCTGGGCATC	<i>NheI</i>
KH46	GCGGATCCGTTTGTGACGACGATGACG	<i>BamHI</i>
KH47	GCGAATTCACGAGCTGAACAGCAGCAAT	<i>EcoRI</i>
KH48	GCAAGCTTCAATATGGCGAGCGATTCTT	<i>HindIII</i>
NMB1796		
KH57	GCGCTAGCAGGCTGCCCACCAAAATACT	<i>NheI</i>
KH58	GCGGATCCCGTTTACCGTACCGTTGCTT	<i>BamHI</i>
KH59	GCGAATTCGTTGAAATCGGACATCTG	<i>EcoRI</i>
KH60	GCAAGCTTGCTATAAACACGCCGGTCAT	<i>HindIII</i>
NMB1963		
KH61	GCGCTAGCTCAAAATACCGATGCCCAAT	<i>NheI</i>
KH62	GCGGATCCGCGGTTCCGACAAAACCTTAC	<i>BamHI</i>
KH63	GCGAATTCGGTAAGCCAAATCCGTCAAA	<i>EcoRI</i>
KH64	GCAAGCTTTTTGCTCGGAAAACCCATAC	<i>HindIII</i>
<i>ssb</i> (NMB1460)		
KH73	GCGCTAGCGCCGATGAGGATGACTTTGT	<i>NheI</i>
KH74	GCGGATCCGCGCTCTATGGTTTCAAAA	<i>BamHI</i>
KH75	GCGAATTCATCCCGAAGTGCGCTATATG	<i>EcoRI</i>
KH76	GCAAGCTTGCCCCGTAAGGTTTAAGGTT	<i>HindIII</i>

*Restriction sites are underlined.

Supplementary Table S4. Primers employed in ComL and SSB recombinant protein construction

Primer	Sequence (5'→3')*	Restriction sites
SAF72	CGTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAAAA AATTCTTTTAACGGTTT	<i>XbaI</i>
SAF73	CGCTCGAGGTGCCAGTAACGCCACCAGGGCATAT	<i>XhoI</i>
HH12	CGCCATGGCCACTCAAGGTACGGTCGATAA	<i>NcoI</i>
HH13	CGCTCGAGCGGGTCGCGGTCGACCAGTCTT	<i>XhoI</i>
HH14	CGCCATGGACCCGAAAGCCAACCGCGAAGCGTAT	<i>NcoI</i>
EH001	TGGGATCCATGTCAATGAACAAAGTCATCCTCAT	<i>BamHI</i>
EH002	CGAAGCTTTCAGAACGGGATATCGTCGTCAATGT	<i>HindIII</i>

*Restriction sites are underlined.

Benam, A. V., Lång, E., Alfsnes, K., Fleckenstein, B., Rowe, A. D., Hovland, E., Ambur, O. H., Frye, S. A. & Tønjum, T. (2011). Structure–function relationships of the competence lipoprotein ComL and SSB in meningococcal DNA transformation. *Microbiology* **157**, 1329–1342.

Supplementary References

- Caugant, D. A., Frøholm, L. O., Bøvre, K., Holten, E., Frasch, C. E., Mocca, L. F., Zollinger, W. D. & Selander, R. K. (1986).** Intercontinental spread of a genetically distinctive complex of clones of *Neisseria meningitidis* causing epidemic disease. *Proc Natl Acad Sci U S A* **83**, 4927–4931. [CrossRef](#) [Medline](#)
- Crowe, B. A., Wall, R. A., Kusecek, B., Neumann, B., Olyhoek, T., Abdillahi, H., Hassan-King, M., Greenwood, B. M., Poolman, J. T. & Achtman, M. (1989).** Clonal and variable properties of *Neisseria meningitidis* isolated from cases and carriers during and after an epidemic in The Gambia, West Africa. *J Infect Dis* **159**, 686–700. [Medline](#)
- Frasch, C. E. & Chapman, S. S. (1972).** Classification of *Neisseria meningitidis* group B into distinct serotypes. I. Serological typing by a microbactericidal method. *Infect Immun* **5**, 98–102. [Medline](#)
- Holten, E. (1979).** Serotypes of *Neisseria meningitidis* isolated from patients in Norway during the first six months of 1978. *J Clin Microbiol* **9**, 186–188. [Medline](#)
- McGuinness, B. T., Clarke, I. N., Lambden, P. R., Barlow, A. K., Poolman, J. T., Jones, D. M. & Heckels, J. E. (1991).** Point mutation in meningococcal por A gene associated with increased endemic disease. *Lancet* **337**, 514–517. [CrossRef](#) [Medline](#)
- Tønjum, T., Freitag, N. E., Namork, E. & Koomey, M. (1995).** Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. *Mol Microbiol* **16**, 451–464. [CrossRef](#) [Medline](#)
- Tønjum, T., Caugant, D. A., Dunham, S. A. & Koomey, M. (1998).** Structure and function of repetitive sequence elements associated with a highly polymorphic domain of the *Neisseria meningitidis* PilQ protein. *Mol Microbiol* **29**, 111–124. [CrossRef](#) [Medline](#)

Characterization of the meningococcal DNA glycosylase Fpg involved in base excision repair

Katrina L Tibballs¹, Ole Herman Ambur², Kristian Alfsnes², Håvard Homberset¹, Stephan A Frye^{1,2}, Tonje Davidsen¹ and Tone Tønjum*^{1,2}

Address: ¹Centre for Molecular Biology and Neuroscience and Institute of Microbiology, University of Oslo, Rikshospitalet, NO-0027 Oslo, Norway and ²Institute of Microbiology, Rikshospitalet, NO-0027 Oslo, Norway

Email: Katrina L Tibballs - k.l.tiballs@medisin.uio.no; Ole Herman Ambur - ole.herman.ambur@rr-research.no; Kristian Alfsnes - kristian.alfsnes@rr-research.no; Håvard Homberset - havard.homberset@rr-research.no; Stephan A Frye - stephan.frye@rr-research.no; Tonje Davidsen - tonje.davidsen@rr-research.no; Tone Tønjum* - tone.tonjum@rr-research.no

* Corresponding author

Published: 9 January 2009

Received: 9 September 2008

BMC Microbiology 2009, 9:7 doi:10.1186/1471-2180-9-7

Accepted: 9 January 2009

This article is available from: <http://www.biomedcentral.com/1471-2180/9/7>

© 2009 Tibballs et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: *Neisseria meningitidis*, the causative agent of meningococcal disease, is exposed to high levels of reactive oxygen species inside its exclusive human host. The DNA glycosylase Fpg of the base excision repair pathway (BER) is a central player in the correction of oxidative DNA damage. This study aimed at characterizing the meningococcal Fpg and its role in DNA repair.

Results: The deduced *N. meningitidis* Fpg amino acid sequence was highly homologous to other Fpg orthologues, with particularly high conservation of functional domains. As for most *N. meningitidis* DNA repair genes, the *fpg* gene contained a DNA uptake sequence mediating efficient transformation of DNA. The recombinant *N. meningitidis* Fpg protein was over-expressed, purified to homogeneity and assessed for enzymatic activity. *N. meningitidis* Fpg was found to remove 2,6-diamino-4-hydroxy-5-formamidopyrimidine (faPy) lesions and 7,8-dihydro-8-oxo-2'-deoxyguanosine (8oxoG) opposite of C, T and G and to a lesser extent opposite of A. Moreover, the *N. meningitidis* *fpg* single mutant was only slightly affected in terms of an increase in the frequency of phase variation as compared to a mismatch repair mutant.

Conclusion: Collectively, these findings show that meningococcal Fpg functions are similar to those of prototype Fpg orthologues in other bacterial species.

Background

Neisseria meningitidis, or the meningococcus (Mc), exclusively colonizes the oro- and nasopharynx of humans. It resides as a commensal in approximately 10% of healthy individuals [1], but may become virulent, disseminating into the bloodstream and crossing the blood-brain barrier

[2]. Mc septicaemia and meningitis are the cause of significant morbidity and mortality worldwide [2].

On the mucosal surface of the upper respiratory tract, Mc is exposed to reactive oxygen species (ROS) produced by the immune system through the oxidative burst and by

endogenous aerobic metabolism, causing damage to many cellular components, including DNA [3]. Oxidative DNA lesions comprise single- and double strand breaks, abasic (apurinic/aprimidinic, or AP) sites, and base damages, among which one of the most common is the oxidation product of guanine, 7,8-dihydro-8-oxo-2'-deoxyguanosine (8oxoG). The mutagenic 8oxoG can mispair with adenine during replication and cause G:C → T:A transversions [4]. 2,6-diamino-4-hydroxy-5-formamidopyrimidine (faPy) is another oxidative modified form of guanine that inhibits DNA synthesis [5].

The base excision DNA repair pathway (BER) is the main defense against the mutagenic and cytotoxic effects of endogenously damaged bases. This enzymatic pathway has been identified in all organisms studied to date [6]. A DNA glycosylase initiates this pathway by cleaving the glycosylic bond between its specific base substrate and the sugar-phosphate backbone, leaving an abasic (AP) site [6]. Many DNA glycosylases also have an inherent AP lyase activity that cleaves the sugar-phosphate backbone at the AP site, which is subsequently repaired by further BER enzymes. In *E. coli*, formamidopyrimidine-DNA glycosylase (Fpg) shows substrate specificity for 8oxoG and faPy lesions, and exhibits AP lyase activity, in successive β- and δ-elimination steps, leaving a single strand break [7].

In *E. coli*, the mutagenic effects of oxidated guanines are prevented by a triplet of enzymes termed the GO system [8]. In GO, Fpg acts together with the DNA glycosylase MutY which removes adenine when mispaired with 8oxoG, and MutT, a nucleotide hydrolase that converts 8oxoGTP to 8oxoGMP, preventing incorporation of oxidized GTPs into the genomic DNA. Mc single *fpg* mutants only elicit a weak mutator phenotype [9], however, *mutY-fpg* double mutants exhibit a much higher increase in

spontaneous mutation frequency than would be expected if *fpg* and *mutY* were involved in unrelated repair mechanisms [9]. This synergistic effect of the two Mc DNA glycosylases confirms their essential role in the repair of oxidative DNA damage and a relationship similar to that in the *E. coli* GO system. *In vivo* Mc Fpg activity has previously been detected in whole cell extracts of clinical isolates by cleavage of 8oxoG opposite C [10], however, the Mc Fpg substrate specificity has not previously been investigated.

In this study, the Mc *fpg* gene was cloned and its gene product over-expressed and purified to homogeneity. Recombinant Mc Fpg was assessed with regard to its enzymatic activity towards recognized Fpg DNA substrates. The Mc MC58 Fpg DNA sequence [11], flanking regions and predicted amino acid sequence was analyzed. Furthermore, sequences of *fpg* homologues and flanking regions in other neisserial species were aligned and examined. Finally, an Mc *fpg* mutant was assessed with regard to phase variation rate and compared to that of the wildtype strain and mismatch repair defective mutants. In essence, the Mc Fpg predicted structure and the activity pattern detected were similar to those of prototype Fpg orthologues in other species.

Methods

Bacterial strains, plasmids, and DNA manipulations

Bacterial strains and plasmids used in this study are listed in Table 1. DNA isolation, PCR amplification and cloning were performed according to standard techniques [12]. The *fpg* gene from Mc strain M1080 was PCR amplified using primers KT1b and KT2b (Table 2). The *fpg*-containing DNA fragment was cloned into the expression vector pET22b, creating plasmid pET22b-*fpg*M1080. *E. coli* ER2566 was used for pET22b-*fpg*M1080 plasmid propaga-

Table 1: Bacterial strains and plasmids used in this study.

Plasmid, strains	Relevant characteristic	Source/Reference
Plasmids		
pET22b	Expression vector, T7 promoter-driven system, His-tag, amp ^R	Novagen
pET22b- <i>fpg</i> M1080	pET22b harbouring <i>fpg</i> from Mc M1080	This study
pARR2107	Contains an Universal Rate Of Switching cassette	[22]
pUD	pARR2107 harbouring a 12-mer DUS	This study
Strains		
<i>Escherichia coli</i>		
ER2566	Expression host with chromosomal copy of the T7 RNA polymerase gene	New England Biolabs
ER2566-pET22b- <i>fpg</i> M1080	ER2566 expressing Mc M1080 <i>fpg</i> from pET22b	This study
<i>Neisseria meningitidis</i>		
M1080	Serogroup B, isolated in the USA in 1970	[45]
Z1099	Serogroup A, isolated in the Philippines in 1968	Dominique A. Caugant
NmZ1099_URO5	Z1099 harbouring a Universal Rate Of Switching cassette	This study
NmZ1099_URO5Δ <i>fpg</i>	Z1099 <i>fpg</i> strain harbouring a Universal Rate Of Switching cassette	This study
NmZ1099_URO5Δ <i>mutS</i>	Z1099 <i>mutS</i> strain harbouring a Universal Rate Of Switching cassette	This study

Table 2: The DNA sequences of primers used in this study.

Oligonucleotide	Sequence (5'-3')*	Application	Source
KT1b	gggaattccatatgcctgaattgccggaagtggaaacg	Cloning	This study
KT2b	cgcgctcagagtttctgacagttcgggcaata	Cloning	This study
TD146	gaagtggaaacgacgttgccgcg	Sequencing	This study
TD147	cgtgccgcgtgccccaaagtcttc	Sequencing	This study
TD160	ctcatacaaagtatgc	Sequencing	This study
TD161	ttgccccaccgtcctgc	Sequencing	This study
TD46	gctgttggaataactggg	Sequencing	This study
TD47	gcatacagataatccgtgc	Sequencing	This study
spcFOR	cccagtggaacataagcctgt	G-tract control, PCR/sequencing	[22]
spcREV	agccgaagtcttccaaaggt	G-tract control, PCR/sequencing	[22]
N248	ggcggcatgacc 8oxog gagggcccatc	DNA substrate Containing 8oxoG lesion	Eurogen
T248	gatgggcctcgggcatgcgcc	DNA substrate, complementary to N248	Eurogen
I393	gatgggcctcagggtcatgcgcc	DNA substrate, complementary to N248	Eurogen
I394	gatgggcctcggggtcatgcgcc	DNA substrate, complementary to N248	Eurogen
I395	gatgggcctctgggtcatgcgcc	DNA substrate, complementary to N248	Eurogen
H7	aacaacaacaaatgccgtcgaacacatccgtctgaaacaacaacaac	Undamaged DNA substrate	This study
H8	gtgtgtgtgttttcagacggcatgtgtgttcagacggcattgtgtgtgt	Undamaged DNA substrate, complementary to H7	This study

* letters in bold represents the DNA lesion or its complementary base in the DNA substrate

tion. DNA sequencing of pET22b-*fpg*M1080 was performed on a Beckman Coulter CEQ8000 Genetic Analyser System (Beckman Instruments, Ca) using an ABI BigDye Terminator v. 3.1 DNA sequencing kit (Applied Biosystems, NC), with primers listed in Table 2.

Bioinformatics analyses and search for signature motifs

An *in silico* search for functional domains and DNA binding motifs was carried out on the deduced amino acid sequence of Mc MC58 Fpg (NMB1295), using the the ExPASy site <http://us.expasy.org/cgi-bin/protscale.pl> and the PROSITE [13] and Pfam databases [14]. The electrostatic charge was calculated using the EMBOSS package <http://emboss.sourceforge.net/>. The DNA sequences of the *fpg* genes and flanking regions as well as the deduced amino acid sequences were aligned and compared in a panel of neisserial species for which the genome sequences were available. The following genome sequences (with accession numbers) were downloaded from Genbank <http://www.ncbi.nlm.nih.gov/>: *Neisseria gonorrhoeae* FA1090 (NC_002946), Mc MC58 serogroup B (NC_03112) [11], Mc Z2491 serogroup A (NC_003116) [15], Mc FAM18 serogroup C (NC_03221) [16] and Mc 053442 serogroup C (NC_010120) [17]. The temporary sequence data for *Neisseria lactamica* ST-640 was obtained from the Pathogen Sequencing Unit at the Sanger Institute <ftp://ftp.sanger.ac.uk/pub/pathogens/nl/>. Access to the genome sequence of Mc 8013 serogroup C was provided by Eduardo Rocha, ABI/Institut Pasteur, Paris, France, with kind permission from Vladimir Pelicic, Necker Hospital, Paris/Imperial College London. Prediction of the Fpg secondary structure was performed based on a blast search <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi> in the JPred program [18]. Protein data bank (PDB) structural

modeling was performed using SMART <http://smart.embl-heidelberg.de/>, Pfam <http://www.sanger.ac.uk/Software/Pfam/>, Phyre <http://www.sbg.bio.ic.ac.uk/phyre/> and PyMol <http://www.pymol.org>.

The Mc *fpg* flanking regions were searched for the presence of transcriptional terminators with the GeSTer genome scanner for terminators (the last version of the program is available at <http://pallab.physics.iisc.ernet.in/~sum05/anil/projects.html> and the TransTermHP program <http://transterm.cbcb.umd.edu/>. Operon predictions were taken from VIMSS [19]. Putative promoters were identified with the transcription promoter predictor available at the Berkeley Drosophila Genome Project http://www.fruit.fly.org/seq_tools/promoter.html and the BPROM predictor of bacterial promoters <http://www.softberry.com/berry.ph.html>. The gene organization of the *fpg* flanking regions in different species was compared using the String program <http://string.embl.de/>.

Purification of the recombinant Mc M1080 Fpg protein

E. coli strain ER2566 over-expressing Mc Fpg encoded by the plasmid pET22b-*fpg*M1080 was grown in LB medium containing ampicillin to a final concentration of 100 µg/ml at 37°C with shaking until OD₆₀₀ was 0.6. The cells were induced with 1 mM isopropyl-D-thiogalactopyranoside and grown for 4 hours. Cells were harvested and washed in phosphate-buffered saline and stored at -70°C. The cells were resuspended in sonication buffer containing 50 mM Na₂HPO₄/NaH₂PO₄, 300 mM NaCl, pH 8.0 and protease inhibitor complete without EDTA (Roche Applied Science, Germany) before lysis by sonication. The cleared lysate was loaded onto a Ni-NTA agarose column

(Qiagen, Germany) and the column washed with wash buffer containing 20 mM imidazole. Bound protein was eluted with a step gradient of 2 column volumes of the elution buffer containing 40, 60, 80, 100, 140, 180, 220 and 250 mM imidazole. Fractions containing purified protein were pooled and dialysed against 25 mM Tris-HCl, pH 7.5, 300 mM NaCl and 10% glycerol.

Assay for base excision of 8oxoG opposite C, A, G or T

Duplex DNA substrates containing a single 8oxoG opposite of C, A, G or T were generated by ^{32}P 5' end-labelling of oligonucleotides, using T4 polynucleotide kinase (New England Biolabs, MA) followed by annealing to a complementary oligonucleotide [20]. The oligonucleotide sequences of the DNA substrates are listed in Table 2. DNA glycosylase reactions were performed by mixing purified protein with 10–50 fmol DNA substrate in a total volume of 10 μl . The enzyme activities were assayed in the reaction buffer previously described [20] and incubated at 37°C for 30 min. *E. coli* Fpg (New England Biolabs, MA) was included as a positive control. Products of the reactions were separated by 20% denaturing PAGE and visualized by phosphorimaging. The assay was performed in triplicate.

Assay for formamidopyrimidine (faPy) DNA glycosylase activity

N-[H3]-*N*-methyl-*N*-nitrosourea (MNU; 1.5 Ci mmol⁻¹) was used to prepare poly(dG-dC) DNA (12,000 dpm mg⁻¹) [21]. DNA glycosylase activity was assayed by mixing purified protein with substrate in a reaction buffer containing 70 mM 3-(*N*-morpholino) propane sulfonic acid, pH 7.5, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 5% glycerol for 30 min at 37°C. Removal of bases was measured in a total reaction volume of 50 μl containing 14 μg of DNA substrate and 500 ng of purified meningococcal protein or 160 U of *E. coli* Fpg (New England Biolabs, MA). The assay was repeated 5 times.

Screening for phase variation by use of a universal rate of switching (UROS) cassette

To promote efficient natural transformation, a 12-mer DNA uptake sequence was inserted into plasmid pARR2107 containing a Universal Rate of Switching (UROS) cassette (kind gift from H. L. Alexander, Emory University School of Medicine, Atlanta, GA) [22], creating plasmid pUD. Mc strain Z1099 (kind gift from D. A. Caugant, Norwegian Institute of Public Health, Oslo, Norway) was transformed with pUD and named NmZ1099_UROS. The *mutS* and *fpg* genes of NmZ1099_UROS were inactivated by insertion of a kanamycin resistance cassette as described by Davidsen *et al.*, 2007 [9] in two separate genetic transformations creating strains NmZ1099_UROS Δ *mutS* and NmZ1099_UROS Δ *fpg*. The mononucleotide tract of 8 G

residues preceding the spectinomycin resistance gene of the UROS cassette was confirmed as an intact 8-mer by PCR and sequencing (by using the primers listed in Table 2) in all three strains before switching frequency/phase variation was assessed. Briefly, Mc strains were grown overnight at 37°C, 5% CO₂, before 10 colonies were resuspended in GC broth. Serial dilutions were plated on GC agar with and without spectinomycin (to a final concentration of 50 mg/l) and incubated overnight. The spectinomycin OFF to ON switching rate was determined by dividing the number of colonies on GC plates containing spectinomycin by the number of colonies on plain GC plates. Phase variation experiments were repeated at least 5 times for each strain. Significance in differences in phase variation frequency was calculated by the Kruskal-Wallis test.

Results and discussion

Fpg is nearly ubiquitous among bacterial species and is highly conserved both within annotated neisserial genome sequences and clinical Mc isolates [10], as well as between evolutionarily distant prokaryotes. We examined the activity and specificity of recombinant Mc Fpg purified to homogeneity towards representative substrates resulting from oxidative DNA damage, 8oxoG and faPy, and detected prototype Fpg glycosylase activity. Previously, we have shown a synergistic effect between the two GO components MutY and Fpg in Mc [9]. Together, these findings emphasize a distinct role for Fpg in the defense against the deleterious effects of reactive oxygen species.

The putative Mc *fpg* open reading frame (ORF) consists of 828 bp and contains a DNA uptake sequence (DUS) (5'-GCCGTCTGAA-3') (Figure 1A). The Mc genome harbours approximately 2000 copies of this highly conserved 10 bp sequence, which is required for efficient transformation [23]. A 12-mer DUS with two additional bp upstream of the core 10 bp repeat element improves the transformation efficiency [24]. The Mc *fpg* gene contains one 11-mer. A single complete DUS or AT-DUS (10-, 11- or 12-mer) may promote the reacquisition of a gene by transformation if it is damaged or deleted and DUS occurs at higher densities in genome maintenance genes than in other house-keeping genes [25].

The organization of the *fpg* flanking region is unique for *Neisseria* species <http://string.embl.de/> (data not shown). Upstream of the *fpg* gene are the hypothetical ORFs NMB1297 and NMB1296 (Figure 1A). NMB1297 is annotated as an ortholog to *mltD* <http://www.ncbi.nlm.nih.gov/COG/>, which encodes a membrane-bound lytic murein transglycosylase of unknown function. NMB1296 shows 30–40% amino acid identity with DNA methyltransferases in a number of bacterial species <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>.

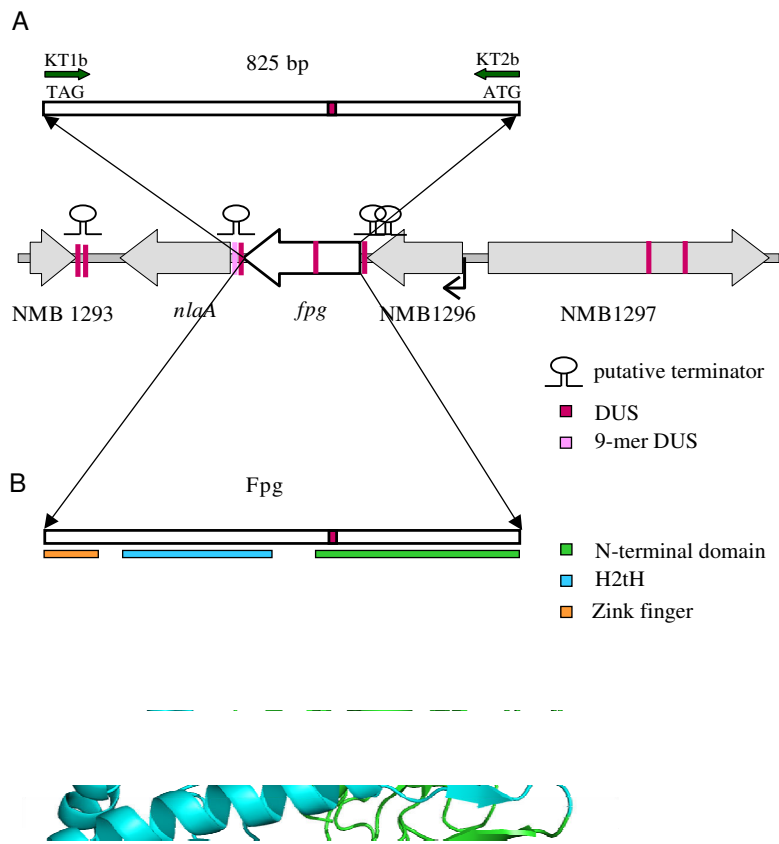


Figure 1
***N. meningitidis* (Mc) *Fpg*.** (A) Physical map of the *Mc fpg* open reading frame and flanking regions. The *fpg* gene contains a DNA uptake sequence (DUS). Primers KT1b and KT2b employed in cloning of the *Mc fpg* gene are depicted. The gene organization of the *Mc fpg* flanking regions is identical in all available neisserial genomes. NMB1296 encodes a hypothetical protein with sequence homology to DNA methyltransferases. A promoter is predicted upstream of NMB1296 (black arrow). The *fpg* and the lysophosphatidic acid acyltransferase *nlaA* genes are putatively co-transcribed [27], although an inverted repeat (containing DUS) associated with transcription termination or attenuation is found downstream of the *fpg* gene. NMB1297 is COG-annotated *mltD* (membrane-bound lytic murein transglycosylase). NMB1293 is a hypothetical protein. The distribution of DUS and degenerate DUS is indicated. (B) Structural modeling of *Mc Fpg* based on *E. coli Fpg* (PDB 1k82) showing the DNA binding motifs helix-two-turn-helix (H2tH) (blue) and zinc finger (orange), as well as the N-terminal domain (green) containing the glycosylase catalytic amino acid residues. Amino acids encoded by DUS are highlighted in purple.

Downstream, *fpg* is flanked by the *nlaA* gene, encoding a lysophosphatidic acid acyltransferase involved in biosynthesis of the glycerophospholipid membrane [26], about 300 bp of non-coding sequence containing two DUS within a predicted terminator, and the opposite oriented hypothetical ORF NMB1293. The NMB1296, *fpg* and *nlaA* genes are all oriented in the same direction and a putative promoter is found upstream of NMB1296 while none are identified between these genes. At the end of NMB1296 a terminator is predicted by TransTermHP. Between *fpg* and *nlaA*, a terminator is predicted by GeSTer. This intrinsic terminator contains a DUS and an imperfect DUS as inverted repeat, a structure found in many putative Mc transcription terminators or attenuators [24]. The VIMSS Operon Prediction suggests co-transcription of *fpg* and NMB1296. However, Swartley and Stephens have evidence by reverse transcriptase PCR that *nlaA* and *fpg* are co-transcribed in Mc strain NMB [27]. In microarray analysis of an MC58 *fpg* mutant compared to wildtype, *nlaA* was the only gene significantly down-regulated at least 1.5 fold, supporting the evidence for co-transcription of these two genes (unpublished data).

The Mc *fpg* open reading frame encodes 276 amino acids containing a predicted N-terminal glycosylase catalytic domain, a helix-two-turn-helix and a C-terminal zinc finger (Figure 1B, additional file 1, Figures S1 and S2). These regions contain long sequences with a positive electrostatic charge, enforcing binding to negatively charged DNA (See additional file 1, Figure S3). Alignment of the deduced Fpg sequence from the genomes of five Mc strains reveals non-synonymous or synonymous substitutions in 5 out of 276 amino acid positions (see additional file 1, Figure S1). The positions showing variation correspond exactly to those found in the *fpg* gene from 11 Mc clinical isolates previously sequenced [10]. An additional 6 amino acids show non-synonymous or synonymous variation when the *N. gonorrhoeae* and *N. lactamica* sequences are included in the comparison. All known functional residues exhibit complete sequence conservation (see additional file 1, Table S1 and Figure S1). Comparison of the neisserial Fpg sequences to those in species where the Fpg crystal structure is solved [28-31] also shows a high degree of conservation, especially in the functional domains and catalytic amino acid residues (see additional file 1, Figure S2). This conservation was confirmed by *in silico* fusion of the crystal structure of *Lactococcus lactis* Fpg with Mc Fpg using the PDB (Figure 1B). Interestingly, the 11-mer DUS sequence encodes amino acids that are not identified as functional residues and is localized in an *fpg* region showing relatively low sequence homology across species boundaries (see additional file 1, Figures S1 and S2).

Fpg has been extensively studied in *E. coli* and is characterized in several other prokaryotes as well [32-34], display-

ing identical substrate specificities. In order to analyze the substrate specificity of Mc Fpg, the gene was over-expressed in *E. coli* and recombinant Mc Fpg protein purified to homogeneity (see additional file 1, Figure S4). Mc Fpg has an apparent size in SDS-PAGE of approximately 30 kDa, corresponding to the molecular weight predicted from the genome deduced amino acid sequence and similar to Fpg of *E. coli* and *L. lactis* [32,33]. The preferred substrates for recognized Fpg proteins are 8oxoG and faPy residues. The ability of recombinant Mc Fpg to remove these lesions was investigated, using *E. coli* Fpg as a positive control. Activity towards C:faPy residues in a ³H-labeled poly(dG-dC) substrate was identified (Table 3). When assessing the 8oxoG excision, the Mc Fpg displayed both DNA glycosylase and AP lyase activity (Figure 2). Equivalent levels of base excision of 8oxoG opposite C, T and G and much lower activity toward 8oxoG when mispaired with A was demonstrated (Figure 2). No activity was detected in the absence of 8oxoG residues (see additional file 1, Figure S5). This discrimination of the base opposite the lesion is in keeping with findings on *E. coli* Fpg [35], although the remaining activity against 8oxoG:A seen in Mc Fpg was not found in the original characterization of substrate specificity in *E. coli*. 8oxoG:C is probably the most important physiological substrate for Mc Fpg, despite the similar levels of nicking observed in 8oxoG:T and 8oxoG:G, as the former is by far the most common substrate *in vivo* in *E. coli* [4]. The removal of 8oxoG from the genome prevents G:C→T:A transversions in *E. coli*, but the mutation rates in single *fpg* mutants are too low in Mc to detect these lesions [9], despite this being the most likely event when 8oxoG is preferentially mis-incorporated with adenine and left unrepaired. Recent studies in *M. smegmatis* have identified an alternative pattern of preferential incorporation of guanine opposite 8oxoG, creating G:C→C:G transversions or A:T→C:G transitions in the absence of Fpg [36]. 8oxoG:G and G:C→C:G transversions can also be found in *E. coli* and *S. pombe*, however, they are rare compared to 8oxoG:A events. In conclusion, these results demonstrate that the protein encoded by the Mc *fpg* gene excises base lesions that are typical substrates of other Fpg orthologues and are consistent with this protein being an Fpg DNA glycosylase.

Mc is a bacterium that seemingly spontaneously produces a plethora of variants upon which selection can act, instead of sensing the environment and changing accordingly [37]. One of the major processes governing genetic changes in *Neisseria* sp. is phase variation. Phase variation is mediated by unstable polynucleotide tracts allowing the gene expression to be switched on or off [37]. Recently, several genome maintenance genes have been shown to modulate phase variation frequencies, including the mismatch repair components *mutS* and *mutL*, the nucleotide excision repair gene *uvrD* and the translesion DNA polymerase *dinB* [38-41]. Since Mc Fpg is able to remove

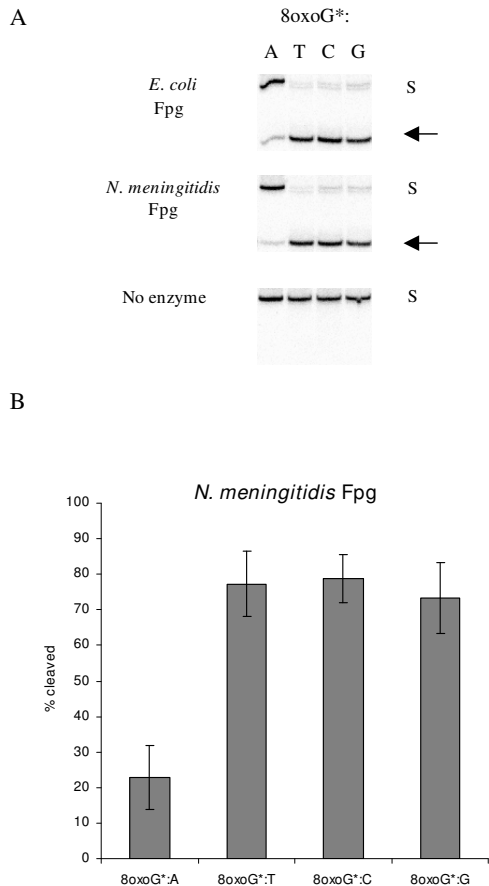


Figure 2
DNA glycosylase activity of *N. meningitidis* (Mc) recombinant Fpg protein. (A) 1 ng of purified Mc Fpg or 0.032 Units of *E. coli* Fpg was incubated with 10–50 fmol of a 24 bp duplex oligodeoxyribonucleotide containing a single 8oxoG residue opposite A, T C or G. Base excision and strand cleavage were analysed by 20% PAGE and phosphorimaging. The arrow indicates the cleaved DNA substrate. * denotes ³²P-labelled strand; S; substrate. (B) Quantification of strand cleavage activity by Mc Fpg. The results represent the average of three independent experiments and error bars indicate the standard deviation of the mean.

oxidized guanines, although in an error-free manner, we wanted to investigate a potential contribution of Mc *fpg* on phase variation of polyG tracts. Mc strains

Table 3: DNA glycosylase activity of *N. meningitidis* (Mc) recombinant Fpg protein.

Substrate	Released bases (fmol)	
	Average	(St. dev.) ^c
<i>N. meningitidis</i> Fpg ^a	75	(± 30)
<i>E. coli</i> Fpg ^b	64	(± 44)
No enzyme	12	(± 4)

^a 500 ng of protein was employed in each reaction

^b 160 Units of protein was employed in each reaction

^c standard deviation of the mean

Removal of formamidopyrimidine (faPy) from [3H]-methyl-faPy-poly(dG-dC) DNA by recombinant Mc and *E. coli* Fpg. The results are given as the average of five independent measurements.

NmZ1099_UROS (Control), NmZ1099_UROSA Δ fpg (Δ fpg) and NmZ1099_UROSA Δ mutS (Δ mutS) were constructed and examined by S12 ribosomal gene switching in a spectinomycin-selection assay (Figure 3). Phase variation was, as previously reported [38-41], significantly increased in the Δ mutS (30-fold) background compared to the wild-type level (** $p < 0.001$). However, the Mc *fpg* mutant exhibited only moderate increase (2-fold) compared to the wild-type level (** $p < 0.001$), and thus MutS exerts a more profound effect on the stability of Mc polyG tracts than Fpg. Likewise, the Mc *fpg* mutant was recently shown to generate only a weak mutator phenotype when assessed for its spontaneous mutation frequency in a rifampicin assay [9]. In conclusion, Fpg is not a major player in modulating Mc mutation frequencies.

Although Mc Fpg displays traits characteristic of the Fpg family of proteins, survival rates of a Mc *fpg* mutant were not affected by exposure to reactive oxygen species [9]. This is in contrast to findings in *M. smegmatis*, where H₂O₂ exposure proved to be lethal to *fpg* null mutants [36], and in the photosynthetic cyanobacteria *S. elongates* where an *fpg*-deficient strain exhibited progressively reduced survival with increasing levels of oxidatively damaging irradiation [42]. Considering the potential importance of oxidative DNA damage in the Mc habitat combined with the vulnerability of a relatively G+C rich genome obtaining such lesions, the explanation for the species discrepancy should be investigated further. The Fpg family of DNA glycosylases also contains endonuclease VIII (Nei) and eukaryotic Nei orthologues. The Nei proteins excise oxidized pyrimidines and may also serve as a backup for removal of 8oxoG in *E. coli* [43], however, no Mc Nei ortholog has been identified [11,15]. On the other hand, the abundant Mc anti-oxidant system provides particularly high protection towards the generation of such DNA lesions [44]. In general, the elucidation of the Mc DNA repair profile is important for understanding the lifestyle

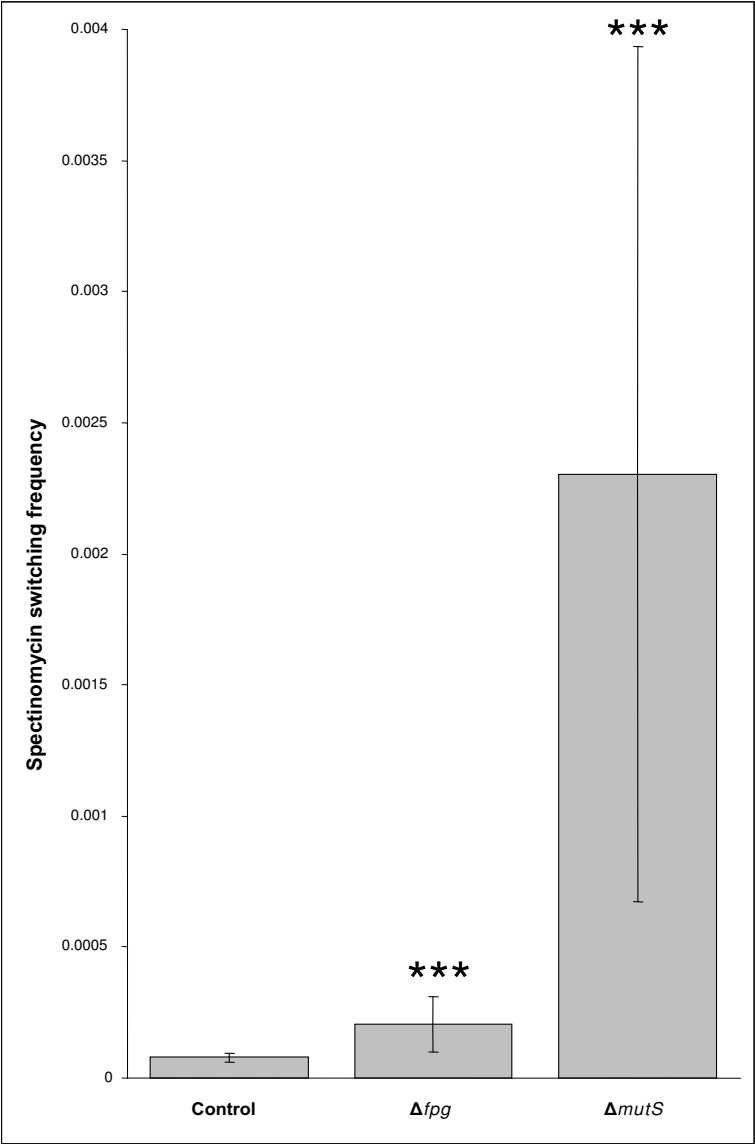


Figure 3
Assessment of meningococcal (Mc) phase variation. Phase variation frequency for Mc strains NmZ1099_UROs (Control), NmZ1099_UROs Δfpg (Δfpg) and NmZ1099_UROs $\Delta mutS$ ($\Delta mutS$) as examined by a spectinomycin assay. The results are given as the median of at least 5 independent measurements. Error bars represent ± 1 quartile. Phase variation is moderately and significantly increased, respectively, in the Mc Δfpg (2-fold) and $\Delta mutS$ (30-fold) background compared to the wild-type level (***) ($p < 0.001$).

of this important pathogen, commensal and model organism.

Conclusion

Mc *fpg* contains DUS both within its coding sequence and in close proximity to the open reading frame, potentially promoting reacquisition of this gene by transformation if it is damaged or lost. The *fpg* gene may belong to an operon together with a putative DNA methyltransferase and a lysophosphatidic acid acyltransferase, although the reasons for this gene organisation remain obscure. Both the nucleotide and amino acid sequences of neisserial Fpg homologues are highly conserved. In addition, Mc Fpg amino acid sequence shows great conservation across species boundaries in functional domains, and Mc Fpg contains a predicted N-terminal glycosylase catalytic domain, a helix-two-turn-helix and a C-terminal zinc finger. Accordingly, Mc Fpg exhibits DNA glycosylase and AP lyase activities and remove both 8oxoG and faPy lesions. When examining the stability of polyG tracts, MutS was found to modulate mutation frequencies due to phase variation to a much higher extent than Fpg. In conclusion, Mc Fpg predicted structure and activity pattern were found to be similar to those of prototype Fpg orthologues in other species. Together, these findings emphasize a distinct role for Mc Fpg in the defense against the deleterious effects of reactive oxygen species.

Authors' contributions

KLT carried out the molecular genetic studies and analysis of purified protein, performed sequence alignments and drafted the manuscript. OHA constructed pUD, designed the phase variation studies and performed the GeSTer analysis. KA contributed to pUD construction and performed the phase variation studies. HH purified recombinant proteins. SAF participated in the bioinformatic analyses. TD supervised the molecular studies and analysis of purified protein, and assisted in manuscript writing. TT conceived the study, participated in its design and coordination and drafted the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Supplementary Material. contains Table S1 Deduced amino acid sequence of Fpg homologues in *Neisseria*, Figure S1 Deduced amino acid sequence of Fpg homologues in *Neisseria*, Figure S2 Deduced amino acid sequence of Fpg orthologues, Figure S3 Electrostatic charge of meningococcal Fpg, Figure S4 Purified meningococcal Fpg, Figure S5 Meningococcal Fpg activity towards undamaged DNA substrate.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2180-9-7-S1.doc]

Acknowledgements

The Medical Research Curriculum at the University of Oslo is greatly acknowledged for its support to KLT. This work was also supported by grants from the Research Council of Norway.

References

1. Yazdankhah SP, Caugant DA: ***Neisseria meningitidis*: an overview of the carriage state.** *J Med Microbiol* 2004, **53**:821-832.
2. Stephens DS, Greenwood B, Brandtzaeg P: **Epidemic meningitis, meningococcaemia, and *Neisseria meningitidis*.** *Lancet* 2007, **369**:2196-2210.
3. O'Rourke EJ, Chevalier C, Pinto AV, Thiberge JM, Ielpi L, Labigne A, Radicella JP: **Pathogen DNA as target for host-generated oxidative stress: role for repair of bacterial DNA damage in *Helicobacter pylori* colonization.** *Proc Natl Acad Sci USA* 2003, **100**:2789-2794.
4. Cheng KC, Cahill DS, Kasai H, Nishimura S, Loeb LA: **8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G-T and A-C substitutions.** *J Biol Chem* 1992, **267**:166-172.
5. Boiteux S, Laval J: **Imidazole open ring 7-methylguanine: an inhibitor of DNA synthesis.** *Biochem Biophys Res Commun* 1983, **110**:552-558.
6. Bjelland S, Seeberg E: **Mutagenicity, toxicity and repair of DNA base damage induced by oxidation.** *Mutat Res* 2003, **531**:37-80.
7. Bhagwat M, Gerlt JA: **3'- and 5'-strand cleavage reactions catalyzed by the Fpg protein from *Escherichia coli* occur via successive beta- and delta-elimination mechanisms, respectively.** *Biochemistry (Masc)* 1996, **35**:659-665.
8. Michaels ML, Miller JH: **The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine).** *J Bacteriol* 1992, **174**:6321-6325.
9. Davidsen T, Tuven HK, Bjoras M, Rodland EA, Tonjum T: **Genetic interactions of DNA repair pathways in the pathogen *Neisseria meningitidis*.** *J Bacteriol* 2007, **189**:5728-5737.
10. Davidsen T, Amundsen EK, Rodland EA, Tonjum T: **DNA repair profiles of disease-associated isolates of *Neisseria meningitidis*.** *FEMS Immunol Med Microbiol* 2007, **49**:243-251.
11. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, Eisen JA, Ketchum KA, Hood DW, Peden JF, Dodson RJ, et al.: **Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.** *Science* 2000, **287**:1809-1815.
12. Sambrook J, Russell DW: **Molecular cloning: a laboratory manual.** Cold Springs Laboratory Press. Cold Spring Harbor, New York; 2001.
13. Hulo N, Sigrist CJ, Le SV, Langendijk-Genevaux PS, Bordoli L, Gattiker A, De CE, Bucher P, Bairoch A: **Recent improvements to the PROSITE database.** *Nucleic Acids Res* 2004, **32**:D134-D137.
14. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, et al.: **The Pfam protein families database.** *Nucleic Acids Res* 2004, **32**:D138-D141.
15. Parkhill J, Achtman M, James KD, Bentley SD, Churcher C, Klee SR, Morelli G, Basham D, Brown D, Chillingworth T, et al.: **Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.** *Nature* 2000, **404**:502-506.
16. Bentley SD, Vernikos GS, Snyder LA, Churcher C, Arrowsmith C, Chillingworth T, Cronin A, Davis PH, Holroyd NE, Jagels K, et al.: **Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18.** *PLoS Genet* 2007, **3**:e23.
17. Peng J, Yang L, Yang F, Yang J, Yan Y, Nie H, Zhang X, Xiong Z, Jiang Y, Cheng F, et al.: **Characterization of ST-4821 complex, a unique *Neisseria meningitidis* clone.** *Genomics* 2008, **91**:78-87.
18. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ: **JPred: a consensus secondary structure prediction server.** *Bioinformatics* 1998, **14**:892-893.
19. Price MN, Huang KH, Alm EJ, Arkin AP: **A novel method for accurate operon predictions in all sequenced prokaryotes.** *Nucleic Acids Res* 2005, **33**:880-892.
20. Eide L, Bjoras M, Pirovano M, Alseth I, Berdal KG, Seeberg E: **Base excision of oxidative purine and pyrimidine DNA damage in *Saccharomyces cerevisiae* by a DNA glycosylase with sequence similarity to endonuclease III from *Escherichia coli*.** *Proc Natl Acad Sci USA* 1996, **93**:10735-10740.

21. Boiteux S, Belleney J, Roques BP, Laval J: **Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides.** *Nucleic Acids Res* 1984, **12**:5429-5439.
22. Alexander HL, Richardson AR, Stojiljkovic I: **Natural transformation and phase variation modulation in *Neisseria meningitidis*.** *Mol Microbiol* 2004, **52**:771-783.
23. Goodman SD, Scocca JJ: **Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*.** *Proc Natl Acad Sci USA* 1988, **85**:6982-6986.
24. Ambur OH, Frye SA, Tonjum T: **New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators.** *J Bacteriol* 2007, **189**:2077-2085.
25. Davidsen T, Rodland EA, Lagesen K, Seeberg E, Rognes T, Tonjum T: **Biased distribution of DNA uptake sequences towards genome maintenance genes.** *Nucleic Acids Res* 2004, **32**:1050-1058.
26. Swartley JS, Balthazar JT, Coleman J, Shafer WM, Stephens DS: **Membrane glycerophospholipid biosynthesis in *Neisseria meningitidis* and *Neisseria gonorrhoeae*: identification, characterization, and mutagenesis of a lysophosphatidic acid acyltransferase.** *Mol Microbiol* 1995, **18**:401-412.
27. Swartley JS, Stephens DS: **Co-transcription of a homologue of the formamidopyrimidine-DNA glycosylase (*fpg*) and lysophosphatidic acid acyltransferase (*nlaA*) in *Neisseria meningitidis*.** *FEMS Microbiol Lett* 1995, **134**:171-176.
28. Sugahara M, Mikawa T, Kumasaka T, Yamamoto M, Kato R, Fukuyama K, Inoue Y, Kuramitsu S: **Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8.** *EMBO J* 2000, **19**:3857-3869.
29. Serre L, Pereira de JK, Boiteux S, Zelwer C, Castaing B: **Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA.** *EMBO J* 2002, **21**:2854-2865.
30. Gilboa R, Zharkov DO, Golan G, Fernandes AS, Gerchman SE, Matz E, Kycia JH, Grollman AP, Shoham G: **Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA.** *J Biol Chem* 2002, **277**:19811-19816.
31. Fromme JC, Verdine GL: **Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM.** *Nat Struct Biol* 2002, **9**:544-552.
32. Boiteux S, O'Connor TR, Lederer F, Gouyette A, Laval J: **Homogeneous *Escherichia coli* FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites.** *J Biol Chem* 1990, **265**:3916-3922.
33. Duwat P, de OR, Ehrlich SD, Boiteux S: **Repair of oxidative DNA damage in gram-positive bacteria: the *Lactococcus lactis* Fpg protein.** *Microbiology* 1995, **141**(Pt 2):411-417.
34. Senturker S, Bauche C, Laval J, Dizdaroğlu M: **Substrate specificity of *Deinococcus radiodurans* Fpg protein.** *Biochemistry (Mosc)* 1999, **38**:9435-9439.
35. Tchou J, Kasai H, Shibutani S, Chung MH, Laval J, Grollman AP, Nishimura S: **8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity.** *Proc Natl Acad Sci USA* 1991, **88**:4690-4694.
36. Jain R, Kumar P, Varshney U: **A distinct role of formamidopyrimidine DNA glycosylase (MutM) in down-regulation of accumulation of G, C mutations and protection against oxidative stress in mycobacteria.** *DNA Repair (Amst)* 2007, **6**:1774-1785.
37. Moxon ER, Rainey PB, Nowak MA, Lenski RE: **Adaptive evolution of highly mutable loci in pathogenic bacteria.** *Curr Biol* 1994, **4**:24-33.
38. Richardson AR, Stojiljkovic I: **Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*.** *Mol Microbiol* 2001, **40**:645-655.
39. Richardson AR, Yu Z, Popovic T, Stojiljkovic I: **Mutator clones of *Neisseria meningitidis* in epidemic serogroup A disease.** *Proc Natl Acad Sci USA* 2002, **99**:6103-6107.
40. Alexander HL, Rasmussen AW, Stojiljkovic I: **Identification of *Neisseria meningitidis* genetic loci involved in the modulation of phase variation frequencies.** *Infect Immun* 2004, **72**:6743-6747.
41. Martin P, Sun L, Hood DW, Moxon ER: **Involvement of genes of genome maintenance in the regulation of phase variation frequencies in *Neisseria meningitidis*.** *Microbiology* 2004, **150**:3001-3012.
42. Muhlenhoff U: **The FAPY-DNA glycosylase (Fpg) is required for survival of the cyanobacterium *Synechococcus elongatus* under high light irradiance.** *FEMS Microbiol Lett* 2000, **187**:127-132.
43. Blaisdell JO, Hatahet Z, Wallace SS: **A novel role for *Escherichia coli* endonuclease VIII in prevention of spontaneous G→T transversions.** *J Bacteriol* 1999, **181**:6396-6402.
44. Seib KL, Tseng HJ, McEwan AG, Apicella MA, Jennings MP: **Defenses against oxidative stress in *Neisseria gonorrhoeae* and *Neisseria meningitidis*: distinctive systems for different lifestyles.** *J Infect Dis* 2004, **190**:136-147.
45. Frasch CE, Gotschlich EC: **An outer membrane protein of *Neisseria meningitidis* group B responsible for serotype specificity.** *J Exp Med* 1974, **140**:87-104.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:

http://www.biomedcentral.com/info/publishing_adv.asp



BioMed Central

Supplementary Material

Characterization of the meningococcal DNA glycosylase Fpg involved in base excision repair

Katrina L Tibballs¹, Ole Herman Ambur², Kristian Alfsnes², Tonje Davidsen¹, Håvard Homberset¹, Stephan A Frye^{1,2} and Tone Tønjum^{1,2*}

Email: Katrina L Tibballs – k.l.tiballs@medisin.uio.no; Ole Herman Ambur – ole.herman.ambur@rr-research.no; Kristian Alfsnes – kristian.alfsnes@rr-research.no; Tonje Davidsen – tonje.davidsen@rr-research.no; Håvard Homberset – havard.homberset@rr-research.no; Stephan A Frye – stephan.frye@rr-research.no; Tone Tønjum – tone.tonjum@rr-research.no

*Corresponding author.

Table S1	Deduced amino acid sequence of Fpg homologues in <i>Neisseria</i>	p 2
Figure S1	Deduced amino acid sequence of Fpg homologues in <i>Neisseria</i>	p 3
Figure S2	Deduced amino acid sequence of Fpg orthologues	p 5
Figure S3	Electrostatic charge of meningococcal Fpg	p 7
Figure S4	Purified meningococcal Fpg	p 8
Figure S5	Meningococcal Fpg activity on undamaged DNA substrate	p 9
	References	p 10

Table S1: Deduced amino acid sequence of Fpg homologues in *Neisseria*.

Comparison of the deduced amino acid sequences of the Fpg homologue in 7 available *Neisseria* sequences. Essential motifs and catalytic residues, as well as the residues encoded by DNA uptake sequence (DUS) are depicted.

Amino acid	Characteristic	Reference
Pro1, Glu2, Glu5, Ile176, Tyr177, Ser221, Thr222, Leu223, Arg224, Gly230 (putative), Gly233 (putative)	Specific 8oxoG binding and lesion recognition	[1-18]
Met74, Arg114, Phe116	Insertion into DNA helix, filling gap after removal of base	[8,12,13,16-18]
Lys57, His 71, Leu 164, Gly172, Gly174, Asn175, Tyr243, Cys 250, Cys 253, Lys257, Arg263, Cys 270, Cys 273	Binding of damaged strand of DNA	[4,8,10,12,13,16,19-26]
His 95, Lys115	Binding of complementary strand of DNA	[8,12,18]
Arg114	Recognition of base opposite 8oxoG	[8,12,18]
Lys57, Glu138, Lys161, Glu180, Gly230 (putative), Gly233 (putative)	Structural formation essential for glycosylase activity	[8,12,13,15,16,19,20,27]
Lys57, Tyr243	β -elimination	[8,12,13,16,19,20]
Tyr243, Arg263	δ -elimination	[8,12,13,16]

```

      *          20          *          40          *
Nl_ST64 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Ng_FA1090 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Nm_MC58 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Nm_053442 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Nm_Z2491 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Nm_8013 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50
Nm_FAM18 : MPELPEVETTLRGIAPHIEGKTVEAVVLRQLKLRWQINPDLGEILSGRQV : 50

      60          *          80          *          100
Nl_ST64 : LSCGRRRAKYLVIRFQTGILLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100
Ng_FA1090 : LSCGRRRAKYLVIRFQTGILLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100
Nm_MC58 : LSCGRRRAKYLLIRFQTGVLLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100
Nm_053442 : LSCGRRRAKYLVIRFQTGILLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100
Nm_Z2491 : LSCGRRRAKYLVIRFQTGILLIHLGMSGSLRIFTSPDGRIGRSDRHDHVDI : 100
Nm_8013 : LSCGRRRAKYLLIRFQTGVLLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100
Nm_FAM18 : LSCGRRRAKYLVIRFQTGILLIHLGMSGSLRIFTSPDGRIGRPDRHDHVDI : 100

      DUS      *          120          *          140          *
Nl_ST64 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCADYLY : 150
Ng_FA1090 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCTDYLY : 150
Nm_MC58 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCADYLY : 150
Nm_053442 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCTDYLY : 150
Nm_Z2491 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCTDYLY : 150
Nm_8013 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCADYLY : 150
Nm_FAM18 : VFSDGTVMRYRDPPKFGAILWYEGIEEHHPLLEKLGPEPLSEAFCADYLY : 150

      160          *          180          *
Nl_ST64 : ARLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Ng_FA1090 : ARLKAQKRAVKLALMDNTVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Nm_MC58 : ARLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Nm_053442 : VRLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Nm_Z2491 : VRLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Nm_8013 : ARLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200
Nm_FAM18 : VRLKAQKRAVKLALMDNAVVVGVGNIIYANESLFRAGISPHRPNRLKKKE : 200

      H2tH
      *          220          *          240          *
Nl_ST64 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Ng_FA1090 : CAVLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Nm_MC58 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Nm_053442 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Nm_Z2491 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Nm_8013 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250
Nm_FAM18 : CALLVETVKAVLRRAIETGGSTLRDFVDSDGKSGYFQQEYTVYGRHNOPC : 250

      260          *
Nl_ST64 : VRCGGGLVVKETILGQRGTIFYCPNCQK* : 275
Ng_FA1090 : LRCGGGLVVKETILGQRGTIFYCTNCQK* : 275
Nm_MC58 : PRCGGGLVVKETILGQRGTIFYCPNCQK* : 275
Nm_053442 : PCGGGLVVKETILGQRGTIFYCPNCQK* : 275
Nm_Z2491 : PCGGGLVVKETILGQRGTIFYCPNCQK* : 275
Nm_8013 : PCGGGLVVKETILGQRGTIFYCPNCQK* : 275
Nm_FAM18 : PCGGGLVVKETILGQRGTIFYCPNCQK* : 275

      zinc finger

```

Figure S1

Deduced amino acid sequence of Fpg homologues in *Neisseria*.

Comparison of the deduced amino acid sequences of the Fpg homologue in 7 available *Neisseria* sequences. Essential motifs and catalytic residues, as well as the residues encoded by the DNA uptake sequence (DUS) are depicted. Colour coding is according to Table S1. Nl: *Neisseria lactamica*, Nm: *Neisseria meningitidis*, Ng: *Neisseria gonorrhoeae*

```

      *           20           *           40
Nm_MC58 : MFELPEVETTLRGIAFHIEGKIVEAV-VLRQLKLRWQI-NPDLGE : 43
E_coli : MPELPEVETSRRGIEHLVGAHILHA-VVRNGRLRWPV-SEIYR : 43
L_lactis : MFELPEVETVRRELEKRIVGQKIVSIEATYPRMVLTF--EQLKK : 43
Tt_HB8 : VPELPEVETTRRLRLVLVGOTLRQV-VHRDP-ARYR--NTALAE : 41
B_stearoth : MFELPEVETIRILLPLIVGKLTIEDVRIFWPNIIRHPRDSEAFAA : 45

      *           60           *           80           *
Nm_MC58 : ILSGRQVLSGRRKAYLLIRFQTG-VLLIHLGMSGSLRIFTSPSDG : 87
E_coli : -LSDQVLSVQRRAKYLLELPEG-WIIIIHLGMSGSLRILPEELP : 86
L_lactis : ELTGKTIHGISRCKYLIFEIGEKDRLISHLRMEGKYRLASLNVP : 88
Tt_HB8 : ---CRRILEVDRRGKFLFALEGGVELVAHLGMTGGFRLEP---- : 79
B_stearoth : RMICQTVRGLERRGKFLKFLDRD-ALISHLRMEGRYAVASALEP : 89

      100      DUS      *           120           *
Nm_MC58 : RIGRPDRHHDHVDIVFSDGTVMRYRDPRRKFGAILWY-EGIEEHHPL : 131
E_coli : ----PEKHHDHVDLVMSNGKVLRYTDPRRKFGAWLWT-KELEGHN-V : 125
L_lactis : ----MEKHHDHLALKFTDEQLI-YADVRRKFGTWELISTDQVLPYFL : 128
Tt_HB8 : -----TPHTRAAALVL-ECRTLYFHDPRRRKFGRLFGVRRGDYREIPL : 118
B_stearoth : ----LEPHHHTHVVFCEFDGSELRVDRRRKFGTMHVYAKEEADRPP : 130

      140           *           160           *           180
Nm_MC58 : IEKLGPEPLSEAFCADYLYARLKAQKRAVKLALMDNAVVGVGNI : 176
E_coli : LTHLGPEPLSDDFNGEYLHQCAKKKTAIKPWIMDNKLVVGVGNI : 170
L_lactis : KKNIGPEPTYTTFDEQIFREKLQKSTKKIKPFLLLEQTLVAGLGNI : 173
Tt_HB8 : LLRLGPEPLSFAFAFPGFRRGLKESARPLKALLDQRLAAGVGNI : 163
B_stearoth : LAELGPEPLSPAFAVLAERAVKTKRSVKALLDQTVVAGVGNI : 175

      *           *           220
Nm_MC58 : YANESLFRAGLSPHRRANRLKKKECALLIVETVKAVLQRAIETGGS : 221
E_coli : YASLSLFAAGTHPDRLASSLSLAECCELLARVIKAVLLRSIEOGGT : 215
L_lactis : YVDEVWLAKLHPEKVANQLTESSIHLHDSIEIIOKAIKLGGS : 218
Tt_HB8 : YADALFRARLSPFRPARSLTEEARRLYRALREVLAEAVELGGS : 208
B_stearoth : YVDESLEFRAGLTPGRPAASLSSKEIERLHEEMVATIGEAVMKGS : 220

      *           *           240           *           260           *
Nm_MC58 : TLRD--FVDSDCKSQYFQOEYTVYGRHNQPCPRCGGLVVKETLGQ : 264
E_coli : TLKD--FLQSDCKPGYFAQELQVYGRKCEPCRVCGGTPIVATKHAQ : 258
L_lactis : SIRT---YSALDCSTGKMODELRVYGKTGEKCVRCGNEIQIKVAG : 260
Tt_HB8 : TLSLQSYRQPDGLPGGFTQTRHAVYGRGELFCPACGRFVERRVVAG : 253
B_stearoth : TVRT--YVNTQCEASTFQHLLYVYGRQGNPCRCGTPIEKTVVAG : 263

      *           *           280
Nm_MC58 : RGTIFYCFNCOQK--- : 275
E_coli : RATFYCRQCOQK--- : 269
L_lactis : RGTHFCHFCQQK-- : 272
Tt_HB8 : RGTHFCHFCQGEGP : 267
B_stearoth : RGTHYCHRCQR--- : 274

```

H2th

zinc finger

Figure S2

Deduced amino acid sequence of Fpg orthologues.

The amino acid sequence of *N. meningitidis* MC58 Fpg compared to species for which the Fpg crystal structure has been solved. Essential motifs and catalytic residues, as well as the residues encoded by DNA uptake sequence (DUS) are depicted. Colour coding according to Table S1. Nm: *Neisseria meningitidis*; E. coli: *Escherichia coli*; L. lactis: *Lactococcus lactis*; Tt HB8: *Thermus thermophilus*; B. stearoth: *Bacillus stearothermophilus*

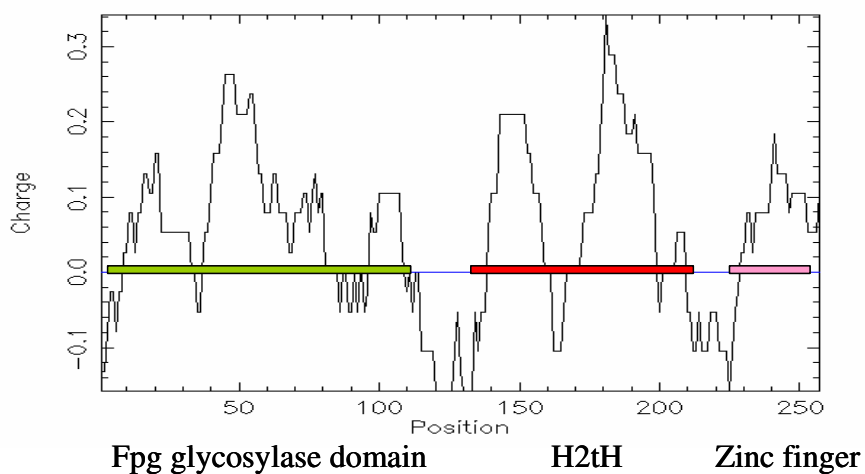


Figure S3

Electrostatic charge of meningococcal Fpg.

Distribution of the predicted electrostatic charge along the meningococcal Fpg amino acid sequence. Functional domains are illustrated; Fpg glycosylase domain (green), helix-2-turn-helix (H2tH) (red), Zinc finger (pink).

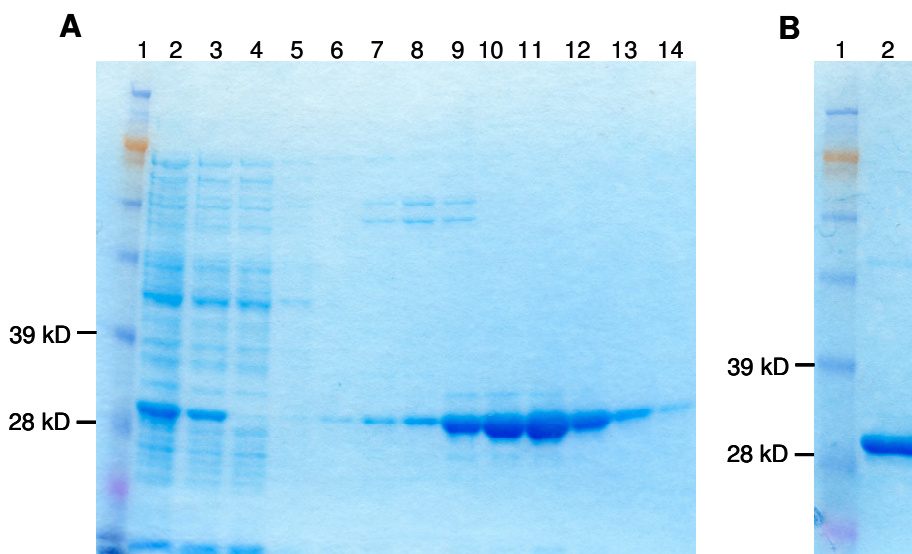


Figure S4

Purified meningococcal Fpg.

(A) Coomassie Brilliant Blue stained SDS-PAGE gel showing the recombinant expression and purification of *N. meningitidis* Fpg protein. Lane 1: SeeBlue® Plus2 Pre-Stained Standard, lane 2: *E. coli* whole cell lysate, lane 3: cleared lysate, lane 4: flow-through, lanes 5 and 6: wash fractions, lanes 7-14: fractions eluted with 40, 60, 80, 100, 140, 180, 220, 250 mM Imidazole, respectively. Mc Fpg has an apparent size in SDS-PAGE of approximately 30 kDa, corresponding to the molecular weight predicted from the genome deduced amino acid sequence. (B) Coomassie Brilliant Blue stained SDS-PAGE gel showing the purified Fpg protein after dialysis. Lane 1: SeeBlue® Plus2 Pre-Stained Standard, lane 2: the pooled elution fraction (100-140 mM Imidazole) that was used in the assays in this study.

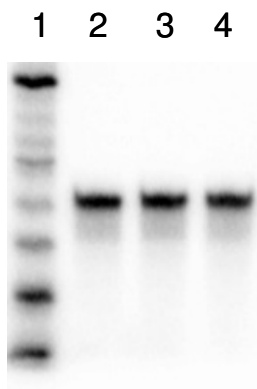


Figure S5

Meningococcal (Mc) Fpg activity on undamaged DNA substrate.

Mc Fpg show no activity towards an undamaged double stranded DNA substrate after 1 hour incubation at 37°C, hence there is no Mc Fpg activity in the absence of 8oxoG residues. Lane 1: ladder, lane 2: double stranded DNA substrate, lane 2: double stranded DNA substrate incubated with 1 ng Mc Fpg, lane 3: double stranded DNA substrate incubated with 50 ng Mc Fpg.

References

1. Pereira de JK, Serre L, Hervouet N, Bouckson-Castaing V, Zelwer C, Castaing B: **Crystallization and preliminary X-ray crystallographic studies of a complex between the *Lactococcus lactis* Fpg DNA-repair enzyme and an abasic site containing DNA.** *Acta Crystallogr D Biol Crystallogr* 2002, **58**:679-682.
2. Saparbaev M, Sidorkina OM, Jurado J, Privezentzev CV, Greenberg MM, Laval J: **Repair of oxidized purines and damaged pyrimidines by *E. coli* Fpg protein: different roles of proline 2 and lysine 57 residues.** *Environ Mol Mutagen* 2002, **39**:10-17.
3. Sidorkina OM, Laval J: **Role of the N-terminal proline residue in the catalytic activities of the *Escherichia coli* Fpg protein.** *J Biol Chem* 2000, **275**:9924-9929.
4. Sugahara M, Mikawa T, Kumasaka T, Yamamoto M, Kato R, Fukuyama K, Inoue Y, Kuramitsu S: **Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8.** *EMBO J* 2000, **19**:3857-3869.
5. Tchou J, Grollman AP: **The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site.** *J Biol Chem* 1995, **270**:11671-11677.
6. Tchou J, Bodepudi V, Shibutani S, Antoshechkin I, Miller J, Grollman AP, Johnson F: **Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA.** *J Biol Chem* 1994, **269**:15318-15324.
7. Zharkov DO, Rieger RA, Iden CR, Grollman AP: **NH₂-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein.** *J Biol Chem* 1997, **272**:5335-5341.
8. Gilboa R, Zharkov DO, Golan G, Fernandes AS, Gerchman SE, Matz E, Kycia JH, Grollman AP, Shoham G: **Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA.** *J Biol Chem* 2002, **277**:19811-19816.
9. Amara P, Serre L, Castaing B, Thomas A: **Insights into the DNA repair process by the formamidopyrimidine-DNA glycosylase investigated by molecular dynamics.** *Protein Sci* 2004, **13**:2009-2021.
10. Amara P, Serre L: **Functional flexibility of *Bacillus stearothermophilus* formamidopyrimidine DNA-glycosylase.** *DNA Repair (Amst)* 2006, **5**:947-958.
11. Pereira de JK, Serre L, Zelwer C, Castaing B: **Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution**

- crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA.** *Nucleic Acids Res* 2005, **33**:5936-5944.
12. Serre L, Pereira de JK, Boiteux S, Zelwer C, Castaing B: **Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA.** *EMBO J* 2002, **21**:2854-2865.
 13. Fromme JC, Verdine GL: **Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM.** *Nat Struct Biol* 2002, **9**:544-552.
 14. Fromme JC, Verdine GL: **DNA lesion recognition by the bacterial repair enzyme MutM.** *J Biol Chem* 2003, **278**:51543-51548.
 15. Lavrukhin OV, Lloyd RS: **Involvement of phylogenetically conserved acidic amino acid residues in catalysis by an oxidative DNA damage enzyme formamidopyrimidine glycosylase.** *Biochemistry (Mosc)* 2000, **39**:15266-15271.
 16. Coste F, Ober M, Carell T, Boiteux S, Zelwer C, Castaing B: **Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase.** *J Biol Chem* 2004, **279**:44074-44083.
 17. Perlow-Poehnelt RA, Zharkov DO, Grollman AP, Broyde S: **Substrate discrimination by formamidopyrimidine-DNA glycosylase: distinguishing interactions within the active site.** *Biochemistry (Mosc)* 2004, **43**:16092-16105.
 18. Zaika EI, Perlow RA, Matz E, Broyde S, Gilboa R, Grollman AP, Zharkov DO: **Substrate discrimination by formamidopyrimidine-DNA glycosylase: a mutational analysis.** *J Biol Chem* 2004, **279**:4849-4861.
 19. Sidorkina OM, Laval J: **Role of lysine-57 in the catalytic activities of *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg protein).** *Nucleic Acids Res* 1998, **26**:5351-5357.
 20. Rogacheva M, Ishchenko A, Saparbaev M, Kuznetsova S, Ogryzko V: **High resolution characterization of formamidopyrimidine-DNA glycosylase interaction with its substrate by chemical cross-linking and mass spectrometry using substrate analogs.** *J Biol Chem* 2006, **281**:32353-32365.
 21. Bhagwat M, Gerlt JA: **3'- and 5'-strand cleavage reactions catalyzed by the Fpg protein from *Escherichia coli* occur via successive beta- and delta-elimination mechanisms, respectively.** *Biochemistry (Mosc)* 1996, **35**:659-665.
 22. Harbut MB, Meador M, Dodson ML, Lloyd RS: **Modulation of the turnover of formamidopyrimidine DNA glycosylase.** *Biochemistry (Mosc)* 2006, **45**:7341-7346.

23. Zharkov DO, Shoham G, Grollman AP: **Structural characterization of the Fpg family of DNA glycosylases.** *DNA Repair (Amst)* 2003, **2**:839-862.
24. Castaing B, Geiger A, Seliger H, Nehls P, Laval J, Zelwer C, Boiteux S: **Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant FPG proteins.** *Nucleic Acids Res* 1993, **21**:2899-2905.
25. O'Connor TR, Graves RJ, de MG, Castaing B, Laval J: **Fpg protein of Escherichia coli is a zinc finger protein whose cysteine residues have a structural and/or functional role.** *J Biol Chem* 1993, **268**:9063-9070.
26. Tchou J, Michaels ML, Miller JH, Grollman AP: **Function of the zinc finger in Escherichia coli Fpg protein.** *J Biol Chem* 1993, **268**:26738-26744.
27. Rabow LE, Kow YW: **Mechanism of action of base release by Escherichia coli Fpg protein: role of lysine 155 in catalysis.** *Biochemistry (Mosc)* 1997, **36**:5084-5096.

